75 JAAR

1920 1995

tijdschrift van het

nederlands elektronica- en radiogenootschap

deel 60 nr. 3 1995

nederlands elektronica- en radiogenootschap

Nederlands Elektronica- en Radiogenootschap
Correspondentie-adres: Postbus 39, 2260 AA Leidschendam.
Gironummer 94746 t.n.v. Penningmeester NERG, Leidschendam.

HET GENOOTSCHAP

Het Nederlands Elektronica- en Radiogenootschap (NERG) is een wetenschappelijke vereniging, gericht op elektronica, telecommunicatie en informatieverwerking.

Het doel van het NERG is om het wetenschappelijk onderzoek op deze gebieden te bevorderen en de verbreiding en toepassing van de verworven kennis te stimuleren.

BESTUUR

LIDMAATSCHAP

Voor lidmaatschap wende men zich via het correspondentie-adres tot de secretaris. Het lidmaatschap van het NERG staat open voor academisch gegradueerden en anderen die door hun kennis en ervaring bij kunnen dragen aan het genootschap. De jaarlijkse contributie bedraagt ingaande 1996 voor gewone leden *f* 75,- en voor junior leden *f* 39,-. Bij automatische incasso wordt *f* 3,- korting verleend.

Gevorderde 1e fase studenten en 2e fase studenten komen in aanmerking voor het junior lidmaatschap en kunnen daartoe contact opnemen met de contactpersoon op hun universiteit.

In bepaalde gevallen kunnen ook andere leden, na overleg met de penningmeester, voor een gereduceerde contributie in aanmerking komen.

De contributie is inclusief abonnement op het Tijdschrift van het NERG en deelname aan vergaderingen, lezingen en excursies.

HET TIJDSCHRIFT

Het tijdschrift verschijnt gemiddeld vijf maal per jaar. Opgenomen worden artikelen op het gebied van de elektronica en de telecommunicatie. Auteurs, die publicatie van hun onderzoek in het tijdschrift overwegen , wordt verzocht vroegtijdig contact op te nemen met de hoofdredacteur of een lid van de redactiecommissie.

Het jaarabonnement van het tijdschrift bedraagt *f* 75,-

REDACTIECOMMISSIE

ISSN 03743853

SUPPORTING GUIDED DISCOVERY WITH COMPUTER SIMULATIONS: THE SMISLE SYSTEM

Wouter R. van Joolingen

Faculty of Educational Science and Technology, University of Twente
Faculty of Philosophy and Social Science, Eindhoven University of Technology



Abstract
Learning with computer simulations is a promising way of giving teachers and students new opportunities for learning. The advantages of computer simulations are discussed, as well as two major problems: how to support learners in finding their way through all the options a simulation offers, and how to support a teacher who wants to create a simulation learning environment? The SMISLE system is discussed as a solution to these two problems.

1 Introduction: learning with computer simulations

In many domains, like mechanics, electronics and chemistry, it is impossible to teach the virtues of a domain solely by providing students with the theory underlying the domain. Practical experience with the domain itself is an essential ingredient for a deep understanding of domain phenomena and theories. Therefore, most study programs in physical and/or technical domain, be it training for first year students or refreshment courses for experienced engineers, will include a practical component in which students perform their own experiments, ranging from predesigned experiments to lab courses in which students do complete, own-initiated experimental studies.

The utility of such lab courses is beyond discussion, but their effectivity can be sub-optimal due to a number of reasons:

- lab availability can be problematic: in order to perform experiments students must come to a laboratory, where space, equipment and personnel may be in short supply;
- in a laboratory it is usually very time consuming to perform experiments, sometimes, performing many experiments in a specific domain can improve insight, but is impossible due to time constraints;
- experiments can be very expensive, due to needs for expensive equipment and/or expensive materials;
- real-life experiments can be very dangerous: in a number of cases doing experiments with real-life equipment can be very dangerous, for the student, for other people and for the environment. For instance, training airline pilots from the beginning in real aeroplanes would be a very dangerous undertaking;
- in experiments, often a number of effects can be observed only indirectly, and in a number of cases the time span for certain phenomena is either very short or very long, implying that it is virtually impossible to observe such phenomena in practice.

Computer simulations provide us with solutions for some or all of these problems. Simulations often are cheaper than real experiments. They are also less constrained in time and place for their use, they can be used to generate large numbers of experiments. Moreover, they are without danger and they allow that the time scale of experiments can be manipulated. Fast processes can be slowed down, slow processes be speeded up. Also simulations can invoke new kinds of visual representations, providing students with new representations of the domain.

Though simulations can never completely replace practical experience, they can have a lot of added value when used in addition to laboratory experience. Presently, this is recognised by many educators, illustrated by the fact that of all programs of computer assisted instruction that are used in higher education in the Netherlands more than 50 % is based on simulations (De Jong, et. al, 1992).

In addition to the more practical features mentioned above, simulations also have a great educational advantage, because they allow for new ways of learning and teaching. The main type of learning that is supported using computer simulation is *discovery learning*. In discovery learning, students are not told directly what the rules in the domain are, but they are offered a situation in which these rules can be discovered.

The idea behind discovery learning is that the resulting knowledge is of a better quality than knowledge that is a result of traditional teaching. In discovery learning knowledge is *constructed* by the students themselves, rather than *transferred* from the teacher to the student. Knowledge constructed by students themselves is deeper rooted and more connected to existing knowledge, because in constructing knowledge one always starts with ones existing knowledge base.

Simulations offer ample opportunities for creating discovery learning situations. With simulations, students can explore the underlying model by designing and performing experiments, analyse the data from these experiments and formulate hypotheses about the model. The simulation provides the students with feedback on their experiments.

In principle such a process for self directed discovery learning can yield good results in terms of the knowledge that the student can acquire, however, in practice, the results often are disappointing. Students often don't know how to proceed in the very open situation offered by a simulation environment. Quite often, they don't know how to do experiments, or which experiment to perform. Also they often don't know how to state hypotheses about the domain. For instance, in a study by Van Joolingen and De Jong (1991), only 40% of the "hypotheses" stated by students actually were hypotheses, in the sense that they were statements about a relation between two or more variables in the domain.

These problems that students have indicate that there is a need for extra support for the student around the computer simulation to keep the advantages of simulations and discovery learning, but repair the problems associated with it. The form of learning in which students are supported is called *guided discovery learning*.

Another problem with discovery learning with simulations is the availability of material. Teachers often are not equipped to design and create simulations

and the support around it by themselves, because they lack both the knowledge and the tools to create such computer programs.
The current article presents a solution to these two problems: the SMISLE system. SMISLE is a toolkit that allows its user to create computer simulations and instructional support for the student around it. The emphasis of SMISLE is the support for discovery learning by means of a number of *instructional measures*, tasks and pieces of information that are given to the student.

The following sections present SMISLE from two viewpoints: the learner, in casu a student of mechanical engineering, and the author, the teacher who has created this simulation environment with the SMISLE toolkit. The simulation discussed is the system SETCOM, in the domain of oscillations, students learn about the characteristics of oscillatory motion with the help of four simulations in increasing order of difficulty. This facility of *model progression* in which a domain is gradually unfolded to the student is one of the instructional measures present in SMISLE.

2. A SMISLE learning scenario

2.1 Simulation

The central part of every SMISLE application is the simulation window. This is the window where the interface to the model is shown and where learners can enter changes to variables and observe output in the form of graphs, values, animations, and the like. Figure 1 shows the simulation window of SETCOM for the first level of model progression. At this first model progression level the learner will see a simple harmonic model, without friction and without external force. The only two input variables that can be changed are the spring constant and the mass. The dynamic output variables displacement and velocity are depicted in a (scrolling) graph and are also available as numbers. The static output variables 'frequency of the motion' and the 'roots of the characteristic equation' are presented in a table. Input variables can be changed while the simulation is running, and effects are immediately visible. The image of the mass suspended from a spring is an animation that follows the simulation.

2.2 Assignments and other means of instructional support

One of the support measures in SMISLE applications is that learners can ask for assignments. Assignments are meant to support the learner in regulating the learning process and intend to prevent the phenomenon of 'floundering' (Goodyear et. al, 1991). SMISLE offers the possibility for creating different types of assignments of which the most important ones are investigation assignments, in which the learner is asked to investigate a specified relation in the model, specification assignments, in which the learner predicts the values of certain variables and optimisation assignments, in which the learner manipulates input variables in such a way that a certain goal is reached.

Figure 2 shows an assignment selection window. The learner can scroll through a list of available assignments and select one. In this case an assignment was selected that asked the learner to investigate the relation between the mass and the eigenfrequency of the system. For investigation assignments the learner should first go to the simulation window and play around until an idea has been formed. By clicking the answer button from Figure 2, an answer window pops up, where, in this case a number of alternatives is presented. The learner selects the alternative that she thinks true and will receive feedback. This feedback can be of any kind: a new assignment, an explanation, an animation etc. The answering modes and the type of feedback differs according to the type of assignment that is selected.

Apart from assignments, a typical SMISLE learning environment, including SETCOM, offers more kinds of instructional support to the learner. Facilities

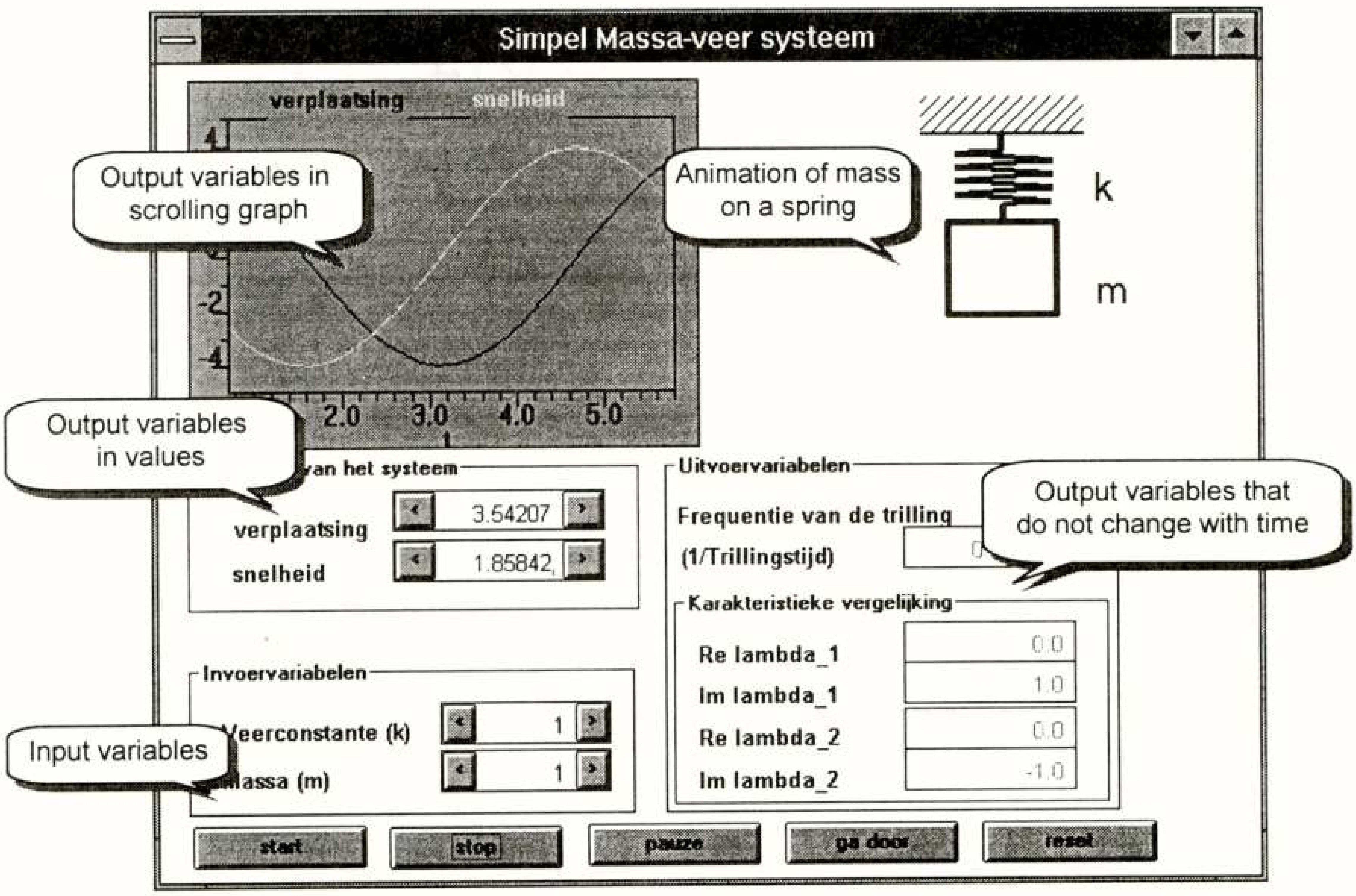


Figure 1 Simulation window at model progression level 1

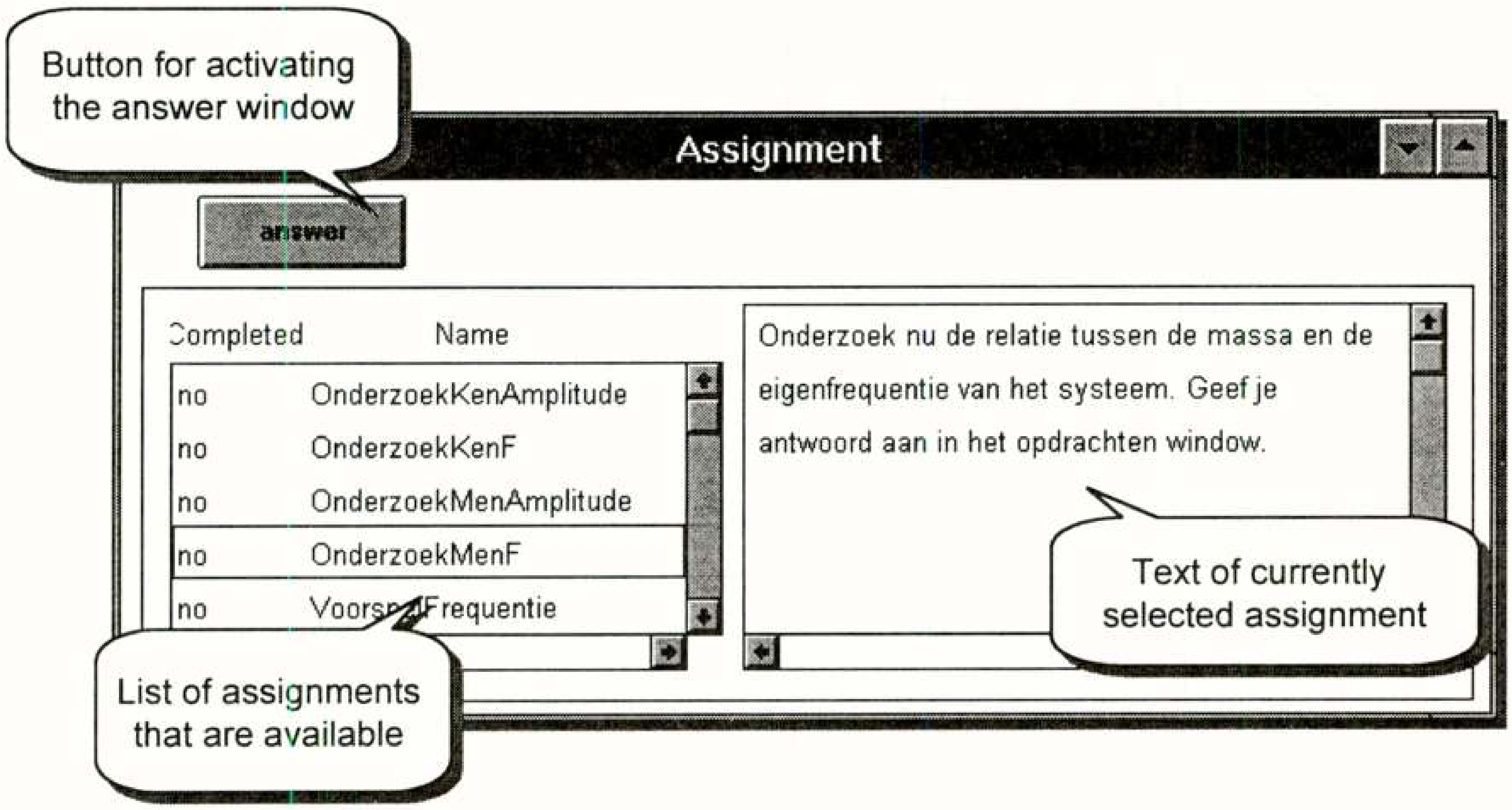


Figure 2 An investigation assignment

are available for explanations on elements of the simulations, model progression, and hypothesis scractchpads, which are notebooks on which the learner can state ideas about the simulation.

A learner who works with a SMISLE learning environment can experiment with the simulation and invoke instructional measures. Behind the scenes, the learning environment keeps track of the status of each instructional measure, for instance of assignments if they are completed. When such a status changes, the learning environment can change the number of available instructional measures for the learner, for instance, when an assignment is completed, new assignments may be added to the list from which the learner selects. In such a way, navigating through the learning environment is a combined activity of both the learner and the learning environments: the learner selects from the set of available instructional measures, and the environment determines the options the learner has to choose from.

3. Authoring a SMISLE environment

An author who wants to create a simulation-based learning environment of oscillatory motion is faced with three tasks:

- creating one or more models of the domain;
- creating a learner interface;
- creating instructional measures for supporting the learner in learning with this simulation.

In the current section the second and third task will be described. For creating the models of the domain, SMISLE offers a modelling tool, which will not be described here. In De Jong et al. (1994) a description of this tool is given, as well as a more elaborate description of all SMISLE tools. In the following subsections it has been assumed that the author has already created a model of the domain.

3.1. Creating a learner interface

For creating a learner interface to a simulation SMISLE offers the author an interface tool. This tool consists of a library of interface elements and a tool for placing the elements in a window and attaching them to variables in the model.

The author has to decide how to represent the model on the screen: just numbers, graphs, static graphics, animation, or a combination of these elements. In the example the author chooses for the latter: graphs, numbers and animation are all present on the screen. The general lay-out of the screen is designed as the author "paints" the interface using elements from the interface library. Elements in the library include numerical controls, sliders, gauges, thermometers, graphs static images and animations. The author selects an element, and uses a *properties editor* to link the interface element to a variable in the simulation. The same editor is used to specify various attributes of the element selected.

During the editing process of the interface, the author can always switch to "learner mode" directly for test purposes. As soon as the author selects "test", the simulation will run just like it will look during a learner session. This facilitates a quick development cycle for both the interface and the underlying model.

3.2. The instructional model

When (the first version of) the simulation and the interface have been created, the author can start with the task which is central to the process of creating a *learning environment*, the creation of the instructional model. From the authors point of view, the instructional model looks like a collection of *instructional measures*, each of which containing a small bit of instructional support, interconnected by mutual dependencies.

The author first has to decide on the nature of the instructional measures to include in the learning environment. This choice is, of course, very situation dependent. In making this choice the author can call upon *author advice* included in SMISLE. This is a hypertext system, containing background information on exploratory learning, information on the instructional measures present in SMISLE, their function and tips for use, and a small expert system, which generates suggestions, based on domain and learner characteristics, entered by the author.

Each instructional measure is created and edited separately, and can then be embedded into the structure of the instructional model. Creating an instructional measure is straightforward: the author selects a template from

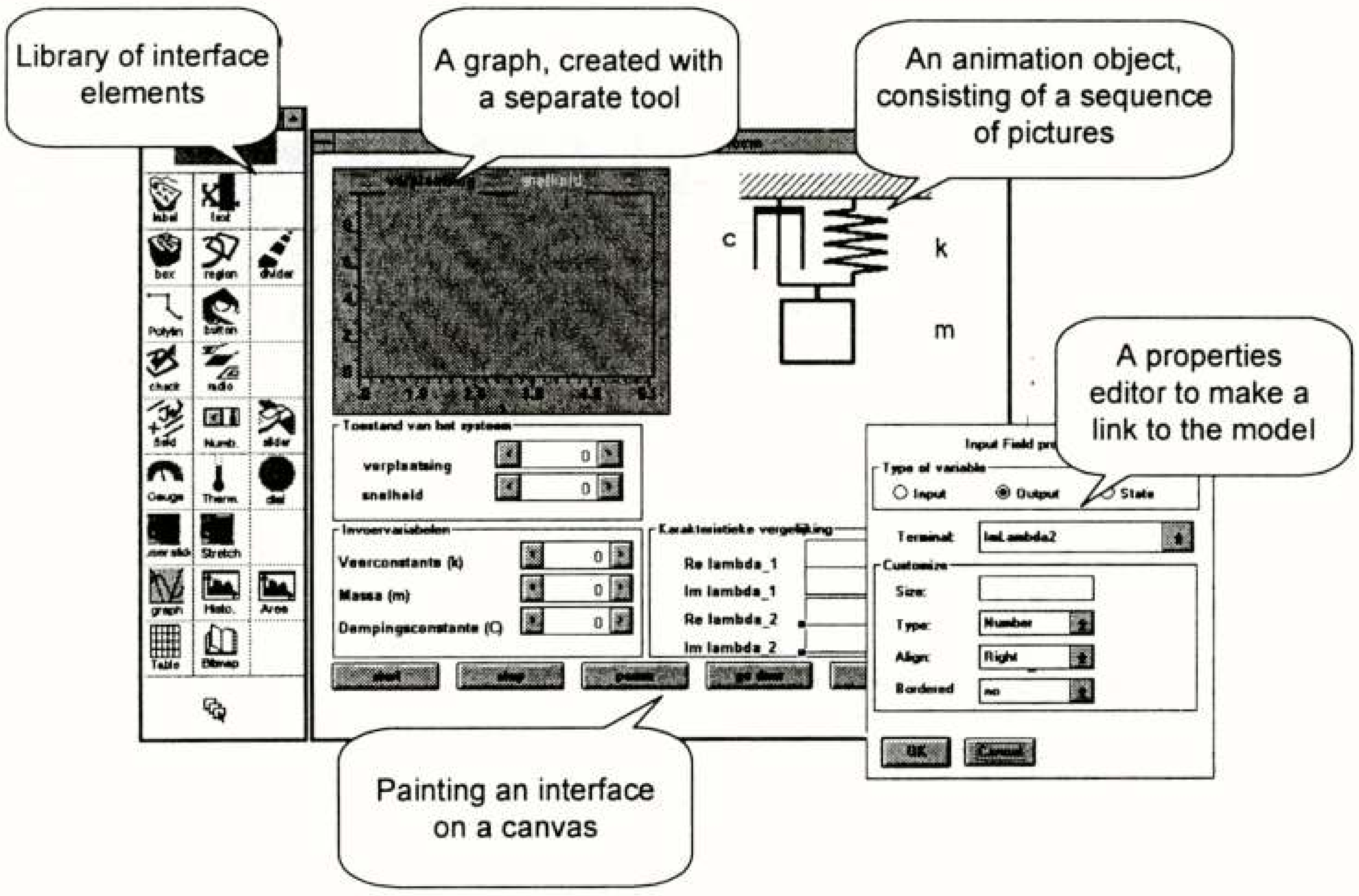


Figure 3 Editing the interface for a model created with SMISLE

the library and then opens a building block editor on it to fill in the details of the particular situation. Figure 4 displays a building block editor for an optimisation assignment. The goal of such an assignment is that the learner finds some optimum values for certain output variables by modifying the inputs of the model. In the case of the system for oscillatory motion, the author includes these assignments because they require the learner to understand the relations between input and output really well. The optimisation assignment is created by setting a goal and constraints. Both are checked during run time: if the goal ("target state") is reached, the assignment is successfully completed; if a constraint is broken, the learner is sent a message and may try again.

The final task for the author is the creation of a set of rules, which determine when a specific instructional measure is activated. The author does so by setting *enabling conditions* for each instructional measure with a dedicated editor. An enabling condition is formulated in terms of other instructional measures, for instance: "assignment #54 is enabled when assignment #33 is completed".

4.Conclusion

In the current paper the advantages of and problems with learning with computer simulations have been discussed. The SMISLE system as a solution for both the problems of the learners as the teachers (authors) of simulation-based tutoring has been described. SMISLE is currently used in a number of locations, and at this moment five simulation-based learning environments have been created with it in various domains: SETCOM, in the domain of oscillations; Collision, in the domain of

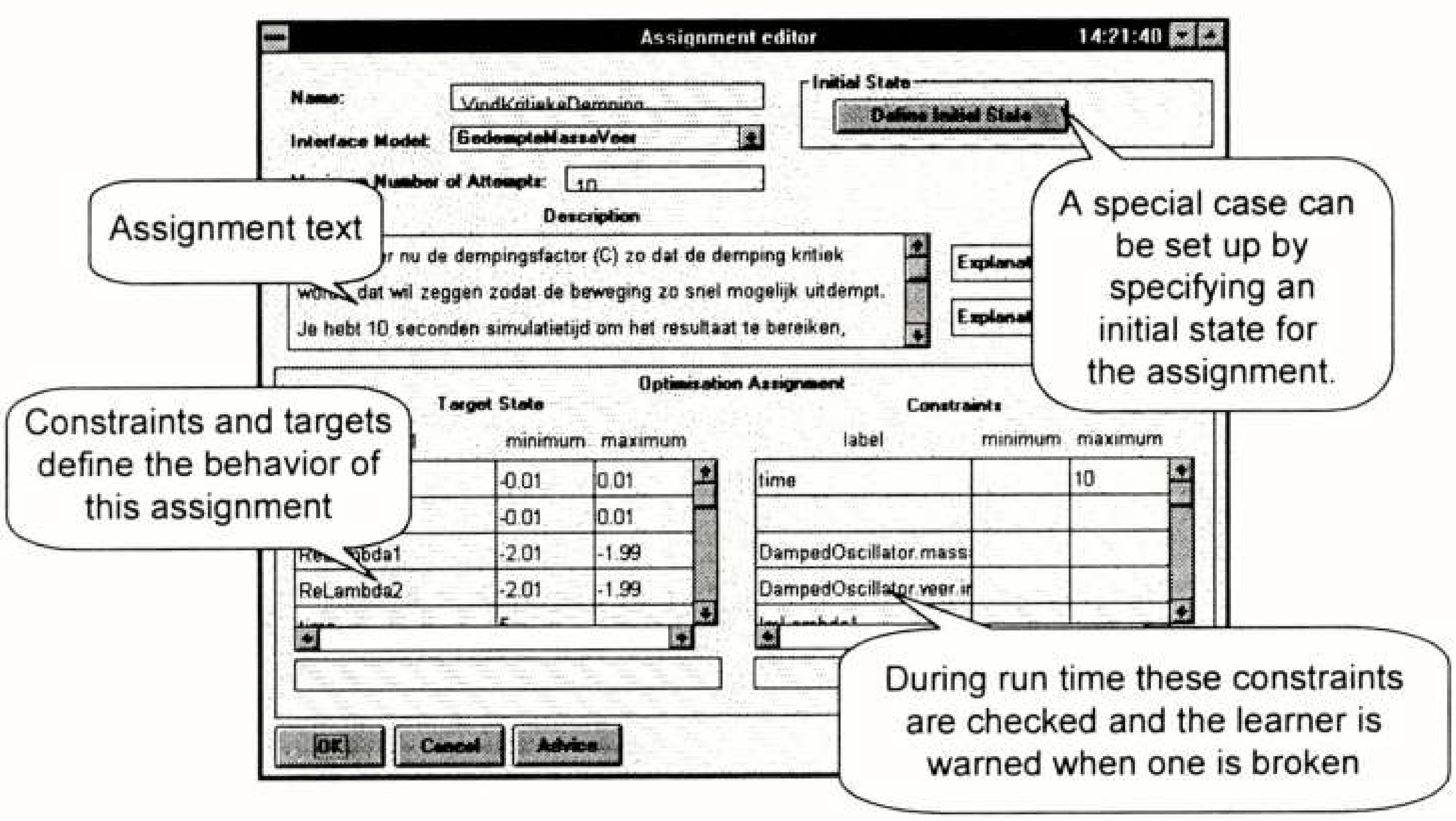


Figure 4 A building block editor to tailor an assignment from the library to the author's situation

collisions; TeEl on telegraph lines; PAMSTAMP, on the design of stamps for metal sheets in car industry and HPU, on the start-up procedure of a hydrogen purification unit in a ethylene plant. These learning environments are currently evaluated with students.

5.Acknowledgement

The SMISLE project is partially funded by the European Union under contract D2007. The author wishes to thank his colleagues in the SMISLE project at the various institutions participating in the project: Framentec-Cognitech, ESI (France), Marconi Simulation (UK), University of Amsterdam and University of Twente (Netherlands), IPN (Germany) and University of Murcia (Spain).

6.References

Goodyear, P., Njoo, M.K.H., & Hijne, H., Berkum, J. van (1991). Learning processes, learner attributes and simulations. *Education and Computing, 6,* 263-304.

Jong, T. de, Andel, J. van, Leiblum, M., & Mirande, M. (1992). Computer assisted learning in higher education in the Netherlands, a review of findings. *Computers and Education, 19,* 381-386.

Jong, T. de, Joolingen, W.R. van, Scott, D., Hoog, R. de, Lapied, L., Valent, R. (1994). SMISLE,: System for Multimedia Integrated Simulation Learning Environments. In: T. de Jong, & L Sarti, (Eds.) *Design and Production of multimedia and simulation-based learning material* (pp. 133-167). Dordrecht: Academic Publishers.

Voordracht gehouden tijdens de 427e werkvergadering

NEW RESULTS ON A FUNDAMENTAL PROBLEM IN NETWORK INFORMATION THEORY

prof. dr. ir. J. P. M. Schalkwijk, ir. H. B. Meeuwissen, and drs. A. H. A. Bloemen
Group of Information and Communication Theory, Dept. EE, Eindhoven University of Technology.

Abstract
This paper is concerned with two new results on coding strategies for the binary multiplying channel. The binary multiplying channel is a two-way channel that models, for example, a wired-AND connection. The coding strategies are described as a progressive subdivision of the unit square into so-called resolution products. The first part of the paper concerns a new class of constructive coding strategies for the binary multiplying channel that achieve surprisingly high transmission rates. The second part of the paper establishes a new region of achievable rate pairs (R_1,R_2) for the binary multiplying channel that includes the equal rate point $R_1=R_2=0.63072$ bit per transmission. A further substantial improvement of the achievable rate region by unit square division is prohibitively difficult.

I. Introduction

A. Network Information Theory

Network information theory is devoted to the transmission of information in network configurations, i.e. communication situations with more than one information flow. Network information theory concerns various computer networks, like for example local area networks (LANs), larger scale networks like ARPANET or BITNET, or the Internet. These networks have evolved during the past decades. At present, many users all over the world employ computer networks to communicate with other users and to retrieve information. If we consider the present interest in the so-called electronic high-way, then the importance of computer networks can only increase.

Network information theory provides definite answers to maximal achievable transmission rates, and to efficient encoding and decoding techniques in multi-user situations. For example, the theory should provide the tools for reliable communication, reduced complexity, and higher bit rates in computer networks. Traditionally, computer networks are considered to be a collection of independent one-way channels. Nevertheless, information theorists have shown already that higher transmission rates are achievable, and that a significant reduction of coding complexity can be attained by considering the communication in computer networks in a larger context.

Network information theory is far from complete. In spite of the various results obtained by researchers so far, there still are a large number of open problems in multi-way communication situations. In analog with single-user information theory, of which the results oftentimes have been implemented successfully, a more complete understanding of multi-user communication problems should be developed in order to employ the available resources in the most efficient manner. This paper is devoted to one of the fundamental problems in multi-user information theory. The problem is simple, well-defined, and constitutes a bottleneck in a class of multi-user problems.

B. The OSI Model

Computer networks are often described with the help of the so-called open systems interconnection (OSI) model. The coding strategies of interest here are simply designed to guarantee error-free data transmissions between two terminals in the physical and data link layer at of the network. Both noiseless data compression techniques and cryptographic techniques can still be applied at each network host. The most surprising fact that results from the coding techniques that are discussed in this paper is that collisions do not by definition require retransmissions like in most contemporary networks.

II. Statement of the Problem

Network information theory is defined by Shannon's [1] 1961 paper on two-way channels (TWC's). The general memoryless TWC is depicted in Fig. 1. Of course, the first channel input letter $X_{t,1}$ at terminal t, $t=1,2$, of the channel input sequence $X_t=(X_{t,1},X_{t,2},\ldots,X_{t,n})$ is dependent on the message θ_t only. The i-th, $2\leq i\leq n$, channel input letter is also based on the previous channel outputs $(Y_1,Y_2,\ldots,Y_{i-1})$. The decoder at terminal t estimates message θ_{3-t} from both the channel output sequence $\mathbf{Y}_t=(Y_{t,1},Y_{t,2},\ldots,Y_{t,n})$, and from the local message θ_t. Both the channel input letters and the channel output letters are taken from alphabets of finite cardinality. The memoryless property of the channel means that two successive channel operations are independent.

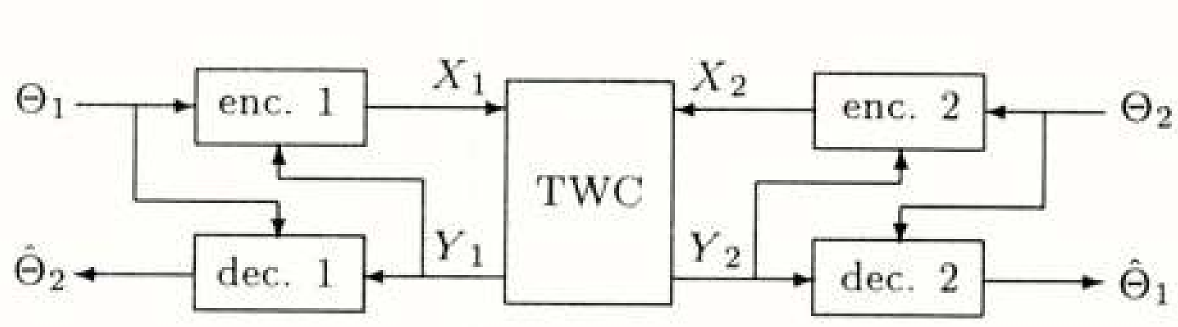


Fig. 1. The general memoryless two-way channel.

Shannon [1] derived single letter inner and outer bound regions to the capacity region of the general memoryless TWC. As a result, the largest region of achievable rate pairs of TWC's with coinciding inner and outer bounds is known. However, D. Blackwell [1] showed that Shannon's inner and outer bound regions are not the same for the binary multiplying channel (BMC), defined by $Y_1=Y_2=Y=X_1X_2$, where $X_1,X_2\in\{0,1\}$. In fact, for symmetric $R_1=R_2$ operation, the Shannon inner bound region satisfies $R_1=R_2=0.61695$, and the Shannon outer bound region satisfies $R_1=R_2=0.69424$. At the present time, the capacity regions of all TWC's that outperform the Shannon inner bound region are still unknown. However, after the publication of [1], time sharing was proven not to be optimal for the TWC. In fact, Shannon's inner bound region, that in general exceeds the time sharing bound, is based on simultaneous transmissions that are allowed to interfere! Nevertheless, the question of whether the two feedback links back to the encoders allow an extension of the achievable rate region of the general TWC beyond the Shannon inner bound region remained unanswered.

After more than three decades the true capacity region of the general TWC is still unknown. The remainder of this paper is devoted to the

BMC. In [3], [4], [5], and [6] various results on the lower bound of the capacity region of the BMC are obtained, while [7] and [8] yield improvements of the BMC's upper bound. Schalkwijk's [4] 1983 strategy for the BMC achieves a region of rate pairs (R_1,R_2) that satisfies $R_1=R_2=0.63056$ bit per transmission in the case of symmetric operation. In addition, the Hekstra and Willems [8] bound proves that rate pairs (R_1,R_2) satisfying $R_1=R_2>0.64628$ are not achievable for the BMC. As a result, a 0.01572 gap exists between lower and upper bound in the equal rate case. Research focused on closing this gap is, in our perspective, important for several reasons. First, the BMC is the simplest example of a non-trivial TWC, and it seems impossible to solve the general TWC as long as the BMC is not completely understood. Second, it seems likely that methods developed for the BMC can be adapted to solve other TWC's. Third, as remarked in [2, ch. 13], many practical channels are intrinsically TWC's.

First, the BMC occurs in optical fibre communication. For example, let a light pulse correspond to a "0", and let no light pulse correspond to a "1". In addition, if we assume that the sensitivity of both receivers is the same for both a light pulse sent by the own transmitter, and a light pulse sent by the other transmitter, then this channel is the BMC. Second, the various devices in (single-chip) microcomputer systems can be connected to a serial bus by a wired-AND. This so-called I^2C-bus concept requires less wiring and fewer connection pins. Note that the channel between two devices is a binary multiplying channel also. Third, wired-AND connections can also be found in so-called controller area networks. These networks are originally used to reduce the amount of cabling in vehicles. If operation on any of these or comparable channels requires more bandwidth or a higher throughput, then two-way channel coding strategies provide a solution.

The paper is organized as follows. In Section III, the concept of coding strategies is explained in the light of Schalkwijk's coding strategies. In Section IV, the so-called discrete save-up strategies are introduced. This class of constructive coding strategies reveals details on the structure of good coding strategies. In Section V, a new coding strategy, based on both the 1983 strategy, and on the discrete save-up strategies is discussed. This strategy achieves a region of rate pairs (R_1,R_2) that allows $R_1=R_2=0.63072$ bit per transmission in the case of symmetric operation. As a result, the gap is closed to 0.01556 in the equal rate case. The reader is referred to [9] for the technical details that are omitted.

III. The Schalkwijk Strategies

The message θ_t at terminal t, $t=1,2$, can without loss of generality be represented by the midpoint of a subinterval $[a_t,b_t)$, $0\leq a_t<b_t\leq 1$, of the unit interval $[0,1)$. Thus, message pairs are subrectangles $[a_1,b_1)\times[a_2,b_2)$ of the unit square $[0,1)\times[0,1)$. The length of each subinterval equals the probability of that subinterval, and the area of each subrectangle equals the probability of that subrectangle. This section is restricted to the case of symmetric $R_1=R_2$ operation.

A. The 1982 Strategy

The 1982 strategy is composed of the three basic resolutions depicted in Fig. 2. The first basic resolution is, if $\theta_t\in[0,\alpha_1)$, $0\leq\alpha_1\leq 1$, then send $X_{t,1}=1$, else if $\theta_t\in[\alpha_1,1)$, then send $X_{t,1}=0$. The second basic resolution is, upon receiving $Y=0$, if $\theta_t\in[0,\alpha_1\alpha_2)\cup[\alpha_1,1)$, $0\leq\alpha_2\leq 1$, then send $X_{t,2}=1$, else if $\theta_t\in[\alpha_1\alpha_2,\alpha_1)$, then send $X_{t,2}=0$. The third basic resolution is, upon receiving $Y=01$, if $\theta_t\in[\alpha_1,1)$, then send $X_{t,3}=1$, else if $\theta_t\in[0,\alpha_1\alpha_2)$, then send $X_{t,3}=0$. After receiving $Y=1$, $Y=00$, $Y=010$, or $Y=011$, the coding strategy repeats itself ad infinitum. Thus, a fractal [10] is generated by the three basic resolutions, and an arbitrary real number between 0 and 1 can be transmitted reliably. Note that decoder t needs message θ_t to estimate θ_{3-t} upon receiving $Y=00$ or $Y=010$.

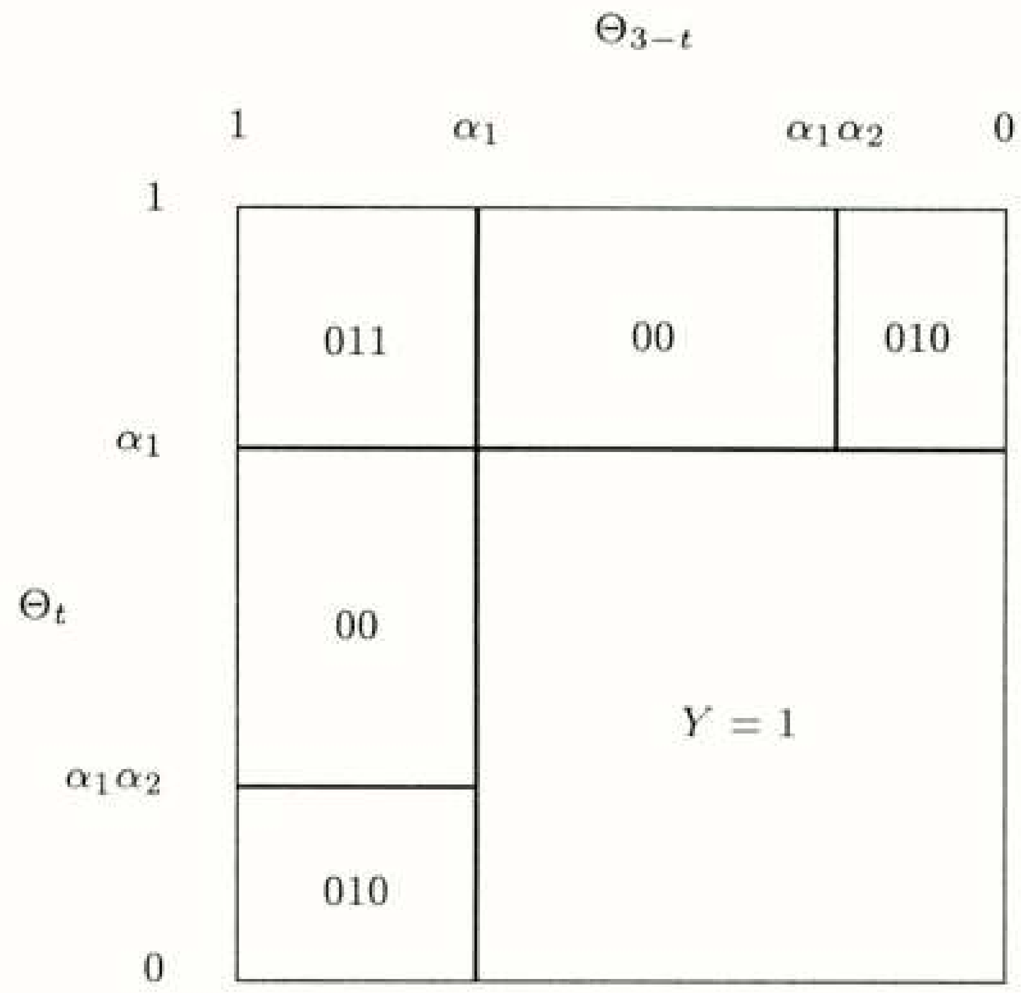


Fig. 2. The three basic resolutions of the 1982 strategy.

From now on, the three basic resolutions will be referred to as the inner, intermediate, and outer bound transmission, respectively. First, the information rate or average mutual information [11] of the inner bound transmission in the direction from terminal t to terminal 3-t is equal to

$$I(\Theta_t;Y|\Theta_{3-t},\ i)=\alpha_1\ h(\alpha_1), \tag{1}$$

where $h(x)=-(1-x)\log_2(1-x)-x\log_2(x)$, $0\leq x\leq 1$, is the binary entropy function, and $0\leq\alpha_1\leq 1$. Second, the information rate of the intermediate transmission satisfies

$$I(\Theta_t;Y|\Theta_{3-t},\ m)=\left(\frac{1}{1+\alpha_1}\right)h(1-\alpha_1+\alpha_1\alpha_2), \tag{2}$$

where $0\leq\alpha_2\leq 1$. Third, the information rate of the outer bound transmission equals

$$I(\Theta_t;Y|\Theta_{3-t},\ o)=\left(\frac{1-\alpha_1+\alpha_1\alpha_2}{1-\alpha_1+2\alpha_1\alpha_2}\right)h\left(\frac{1-\alpha_1}{1-\alpha_1+\alpha_1\alpha_2}\right). \tag{3}$$

The average code word length or probability of the inner bound transmission is equal to $\Pr[i]=1$, the probability of the intermediate transmission equals $\Pr[m]=1-\alpha_1^2$, and the probability of the outer bound transmission is $\Pr[o]=(1-\alpha_1)(1-\alpha_1+2\alpha_1\alpha_2)$. Then, by definition, the overall transmission rate of the 1982 strategy is equal to

$$R=\frac{\Pr[i]\ I(\Theta_t;Y|\Theta_{3-t},\ i)+\Pr[m]\ I(\Theta_t;Y|\Theta_{3-t},\ m)+\Pr[o]\ I(\Theta_t;Y|\Theta_{3-t},\ o)}{\Pr[i]+\Pr[m]+\Pr[o]}. \tag{4}$$

Finally, numerical optimization of (4) yields $\tilde{R}_{1982}$=0.61914 for α_1=0.67571 and α_2=0.29769. The 1982 strategy is constructive. Blahut [2, ch. 13] showed how to make Schalkwijk's 1982 strategy practical with a slight modification only.

B. The 1983 Strategy

The 1983 strategy disposes of the second basic resolution in the 1982 strategy. In other words, the intermediate transmission can be accomplished without affecting the rate of the strategy. The technique involved is called bootstrapping. Thus, the 1983 strategy is generated by inner and outer bound transmissions only. As a result, the overall transmission rate of the 1983 strategy is

$$R=\frac{\Pr[i]\; I(\Theta_t;Y|\Theta_{3-t},\ i)+\Pr[o]\; I(\Theta_t;Y|\Theta_{3-t},\ o)}{\Pr[i]+\Pr[o]}. \qquad (5)$$

Finally, numerical optimization of (5) yields $\tilde{R}_{1983}$=0.63056 for α_1=0.69070 and α_2=0.32060. Thus, bootstrapping can be considered as a technique to enlarge the achievable rate region of a coding strategy. As a side-effect, the improved coding strategy of itself is non-constructive. However, van Overveld [5], [12] proved that the achievable rate region of a coding strategy with bootstrapping remains operationally achievable.

IV. Discrete Coding Strategies

The message θ_t at terminal t, t=1,2, of a discrete coding strategy are drawn according to an independent and uniform distribution from the finite message set {1,2,...,M_t}. As a result, a discrete coding strategy is a progressive subdivision of the $M_1 \times M_2$ rectangle into so-called resolution products.

A. Regular Discrete Coding Strategies

A regular discrete coding strategy subdivides the $M_1 \times M_2$ rectangle into basic 1×1 squares. As an example, consider the Hagelbarger [1] code for $M_1=M_2=2$. If θ_t=1, then send $X_{t,1}$=1, else if θ_t=2, then send $X_{t,1}$=0. Upon receiving Y=0, if θ_t=2, then send $X_{t,2}$=1, else if θ_t=1, then send $X_{t,2}$=0. Thus, the average mutual information is equal to $\log_2(2)$=1 bit, and the average number of transmissions is 7/4. Therefore, the Hagelbarger code achieves $R_1=R_2$=4/7 in excess of the time sharing bound of $R_1=R_2$=1/2. The Schalkwijk [13] code of rate $R_1=R_2=(3/8)\log_2(3)$ for $M_1=M_2=3$ is depicted in Fig. 3. Observe that the Schalkwijk code contains the Hagelbarger code.

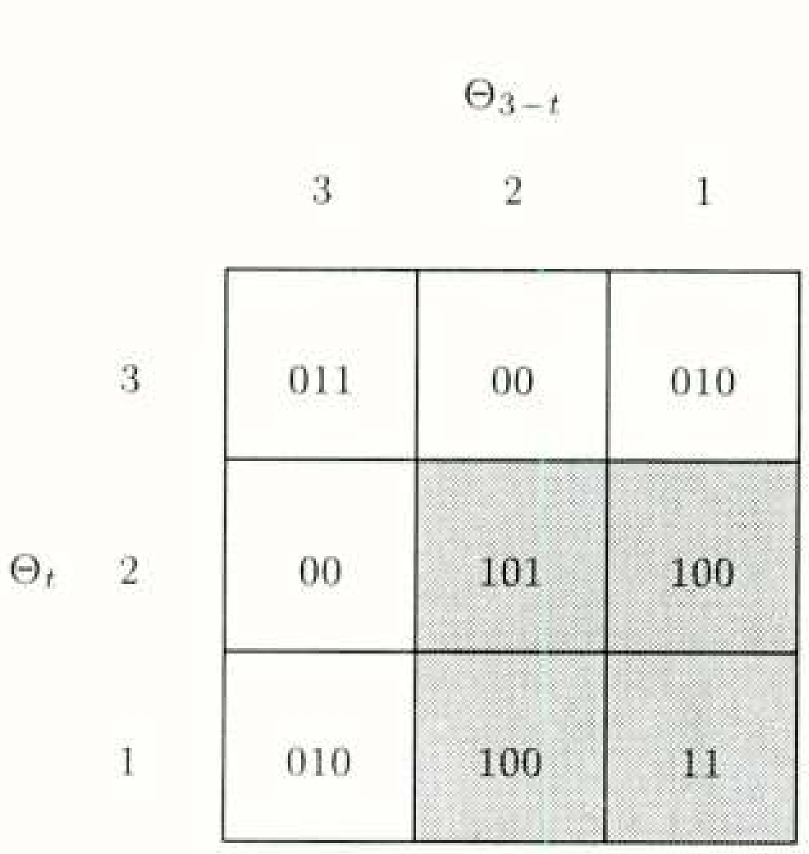


Fig. 3. The Schalkwijk code.

First, note that the lengths of the code words of the strategies in this class of zero-error coding strategies are variable. Second, note that the decoding procedure of a given strategy is surprisingly simple. However, finding the coding strategies is hard. Although van Overveld [12] proved that the capacity region of this class of coding strategies is equal to the true capacity region of the BMC, no regular discrete coding strategies with rate pairs in excess of the Shannon inner bound region are known.

B. Discrete Save-Up Strategies

A discrete save-up strategy subdivides the $M_1 \times M_2$ rectangle into arbitrary rectangular resolution products. As a result, this class of coding strategies avoids low rates on rectangular resolution products. For example, time sharing resolutions that subdivide resolution products of size 1×2 or size 2×1 are no longer completed, since such resolutions affect the overall transmission rate of a coding strategy in a negative way. In fact, some of the information is saved up for later transmission at the rate of the strategy.

The class of discrete save-up strategies is regarded as an extension of Schalkwijk's 1982 strategy on a grid. The subdivisions according to both the 1982 strategy, and the save-up strategies have to satisfy one constraint, i.e. to leave rectangular resolution products only. In addition, the save-up strategies introduce resolution products with all sorts of shapes. In other words, the save-up strategies often consist of more than three basic resolutions. Note also that the basic resolutions of a save-up strategy that subdivides one particular $M_1 \times M_2$ rectangle also can be used to generate a fractal in the unit square by repeating the basic resolutions ad infinitum in all the rectangles. The capacity region of the class of save-up strategies is again equal to the true capacity region of the BMC. An additional advantage of discrete save-up strategies is that they are constructive, just like the 1982 strategy.

However, the problem of finding discrete save-up strategies is also hard. An exhaustive computer search seems infeasible, while a restricted search neither guarantees optimality, nor yields high rates. Up to now, the human mind has been more successful in finding good coding strategies. The discrete save-up strategies are constructed in an environment of computer-aided design of coding strategies, in which the computer only performs the tasks it is most suitable for, i.e. compute rates, compare results, and store relevant data.
This approach yields the following results for symmetric $R_1=R_2=R$ operation in $M \times M$ squares up to size 33. The coding strategies for M=3,4,5,6,7, and 9 are instances of the 1982 strategy on a grid. Thus, a computer optimization on the unit square of the coding strategies for these values of M yields R=0.61914, i.e. the maximal half-sum rate of the 1982 strategy. The coding strategies for M=6, and $M \geq 9$ exceed the $R_1=R_2$=0.61695 rate of the Shannon inner bound. The coding strategies for $M \geq 11$ outperform the $R_1=R_2$=0.61914 rate of the 1982 strategy. For example, the save-up strategy for M=33 achieves a rate of R=0.62786, which is close to the $R_1=R_2$=0.63056 rate of the non-constructive 1983 strategy. Thus, the save-up strategies achieve surprisingly high rates even for low values of M.

To conclude, the save-up strategies uncover new properties on the structure of good coding strategies, since they consist of more than three basic resolutions for most values of M. In the next Section, a new coding strategy is constructed in the unit square by combining an efficient resolution product of the save-up strategies, and an alternative version of the 1983 strategy. As a result, a new achievable rate region for the BMC can be established.

V. The New Coding Strategy

A. An Alternative Version of the 1983 Strategy

The alternative version of the 1983 strategy consists of both inner and outer bound transmissions. The inner bound transmissions are depicted in Fig. 4. Recall that inner bound transmissions are always specified by the parameter α_1.

Three inner bound transmissions generate a resolution product referred to as the outer rim. The first inner bound transmission of probability Pr[i] informs both terminals whether the message pair (θ_1,θ_2) lies in the Y=0 part, or in the Y=1 part. In fact, the first inner bound transmission is a copy of the inner bound transmission of the original 1983 strategy. The second inner bound transmission of probability $\left(1-\alpha_1^2\right)\,\Pr[i]$ subdivides the Y=0 part into both the Y=00 part, and the Y=01 part, such that the quotient of the area of the Y=0 part, and the area of the Y=01 part is equal to $1/\alpha_1^2$. The third inner bound transmission of probability $\left(1-\alpha_1^2\right)^2\,\Pr[i]$ subdivides the Y=00 part into the Y=000 part, and Y=001 part, such that the quotient of the area of the Y=00 part, and the area of the Y=001 part equals $1/\alpha_1^2$. The Y=000 part is the so-called outer rim.

An infinite number of so-called inner rims is generated by inner bound transmissions in the Y=001 part. The ratio between the area of the outer rim, and the area of the first inner rim is $1/\alpha_1^2$. Furthermore, the area of a particular inner rim is equal to the area of the preceding inner rim scaled by the factor α_1^2. As a result, the average code word length of inner bound transmissions in the Y=001 part is equal to

$$\alpha_1^2\left(1-\alpha_1^2\right)^2\,\Pr[i]\sum_{i=0}^{\infty}\left(\alpha_1^2\right)^i=\alpha_1^2\left(1-\alpha_1^2\right)\,\Pr[i]. \tag{6}$$

The subdivision of the Y=0 part is recursively repeated in the Y=01 part, each time after scaling by the factor α_1^2. As a result, the average code word length of inner bound transmissions in the Y=01 part is equal to

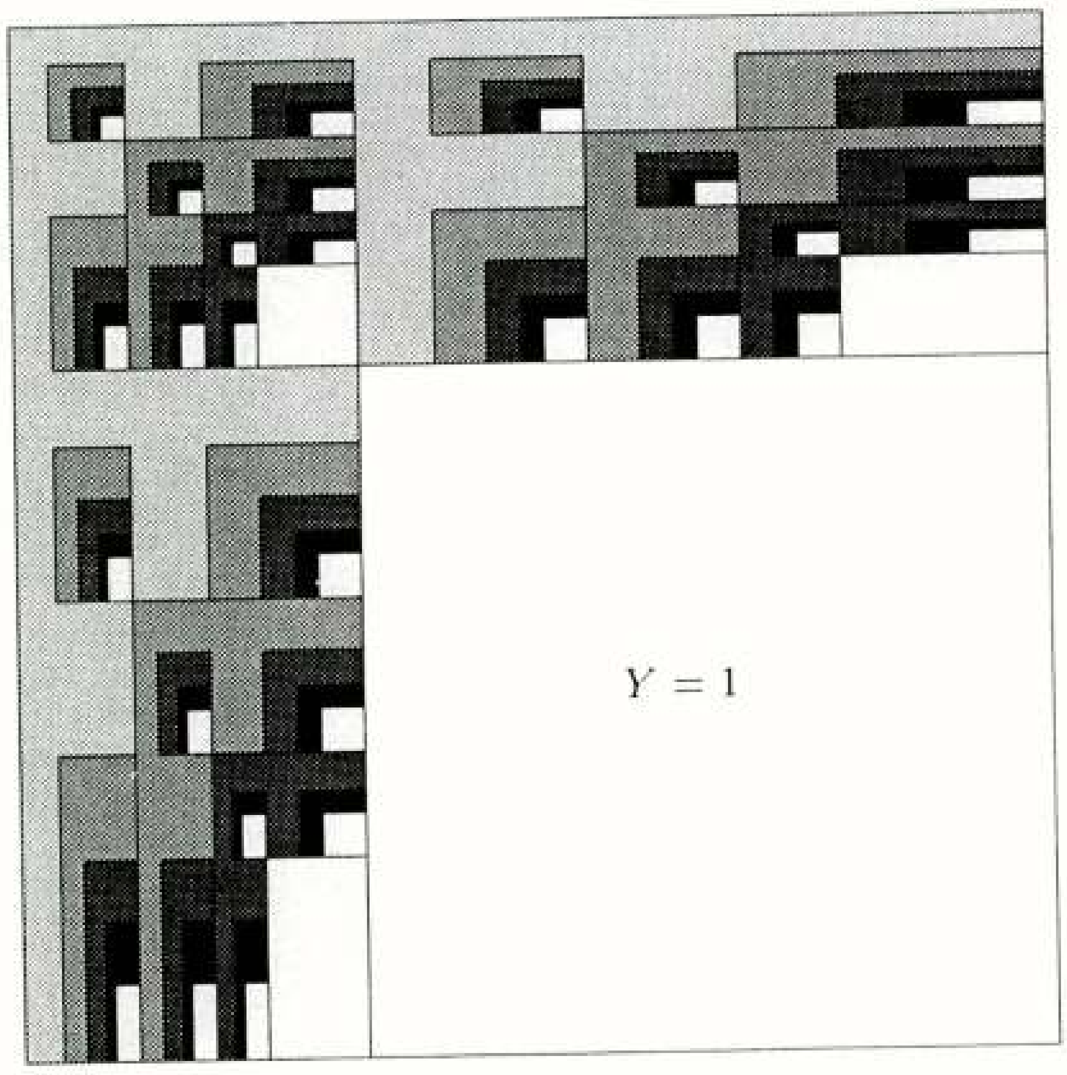


Fig. 4. The inner bound transmissions of the new strategy.

$$\left(\alpha_1^2\left(1-\alpha_1^2\right)\sum_{i=0}^{\infty}\left(\alpha_1^2\right)^i+\alpha_1^2\left(1-\alpha_1^2\right)^2\sum_{i=0}^{\infty}\left(\alpha_1^2\right)^i\sum_{j=0}^{\infty}\left(\alpha_1^2\right)^j\right)\Pr[i]=2\alpha_1^2\,\Pr[i]. \tag{7}$$

In conclusion, the average code word length of the inner bound transmissions in the alternative version of the 1983 strategy sums up to 3Pr[i].

Next, the outer rim is subdivided into rectangular resolution products by several outer bound transmissions. The average code word length of outer bound transmissions in the outer rim equals

$$3\left(1-\alpha_1^2\right)^2\,\Pr[o]. \tag{8}$$

Then, the resolution of the outer rim can be repeated in all the inner rims. As a result, the average code word length of outer bound transmissions in the alternative version of the 1983 strategy is equal to 3Pr[o]. Therefore, the overall transmission rate is equal to

$$R=\frac{3\Pr[i]\;I(\Theta_i;Y|\Theta_{3-i},\,i)+3\Pr[o]\;I(\Theta_i;Y|\Theta_{3-i},\,o)}{3\Pr[i]+3\Pr[o]}. \tag{9}$$

To conclude, the overall transmission rates as given in (5) and (9) are exactly the same. The just described inner and outer bound transmissions are the basic resolutions of the alternative version of the 1983 strategy. Note that these basic resolutions generate a fractal of itself, but, of course, the basic resolutions are repeated in all rectangular resolution products again. Thus, the alternative version of the 1983 strategy is a fractal generated by a fractal.

B. The Loss and the Gain

The new coding strategy starts from the alternative version of the 1983 strategy. Note that the Y=0 part of the unit square contains rims everywhere after the basic inner bound transmissions, and that the rims, according to the alternative version of the 1983 strategy, are resolved by outer bound transmissions only. Now the new coding strategy modifies the resolution of these rims. Note that all the rims can be modified in the same manner, since two distinct rims differ from each other by at most a certain scaling factor. Some of the outer bound transmissions are no longer completed, which results in a loss of both code word length, and average mutual information. In addition, three new transmissions are introduced, which results in a gain of both code word length, and average mutual information.

Let L[*loss*] denote the average code word length of outer bound transmissions that are no longer completed, let L[*gain*] denote the average code word length of the three new transmissions, and let I[*gain*] denote the average mutual information of the three new transmissions, then the rate of the new coding strategy is equal to

$$R=\frac{3\Pr[i]\;I(\Theta_i;Y|\Theta_{3-i},\,i)+\left(3\Pr[o]-\mathrm{L}[loss]\right)\;I(\Theta_i;Y|\Theta_{3-i},\,o)+\mathrm{I}[gain]}{3\Pr[i]+3\Pr[o]-\mathrm{L}[loss]+\mathrm{L}[gain]}. \tag{10}$$

Of course, if the loss overcompensates the gain, then an improvement is obtained. Note that L[*loss*]=0 implies L[*gain*]=0 and I[*gain*]=0, and that substitution in (10) yields (9). Thus, the new coding strategy is at least able to achieve the rate of the 1983 strategy. However, a numerical optimization of (10) with respect to seven parameters yields a rate of

$\bar{R}_{1994}$=0.63072 bit per transmission. To conclude, a substantial improvement of the lower bound to the capacity region of the BMC for symmetric R_1=R_2 operation is obtained. The new lower bound clearly proofs the non-optimality of the 1983 strategy.

The three new transmissions in the outer rim that originate from the discrete save-up strategies are depicted in Fig. 5. The shaded areas are removed by bootstrapping. The three new transmissions replace some of the earlier outer bound transmissions.

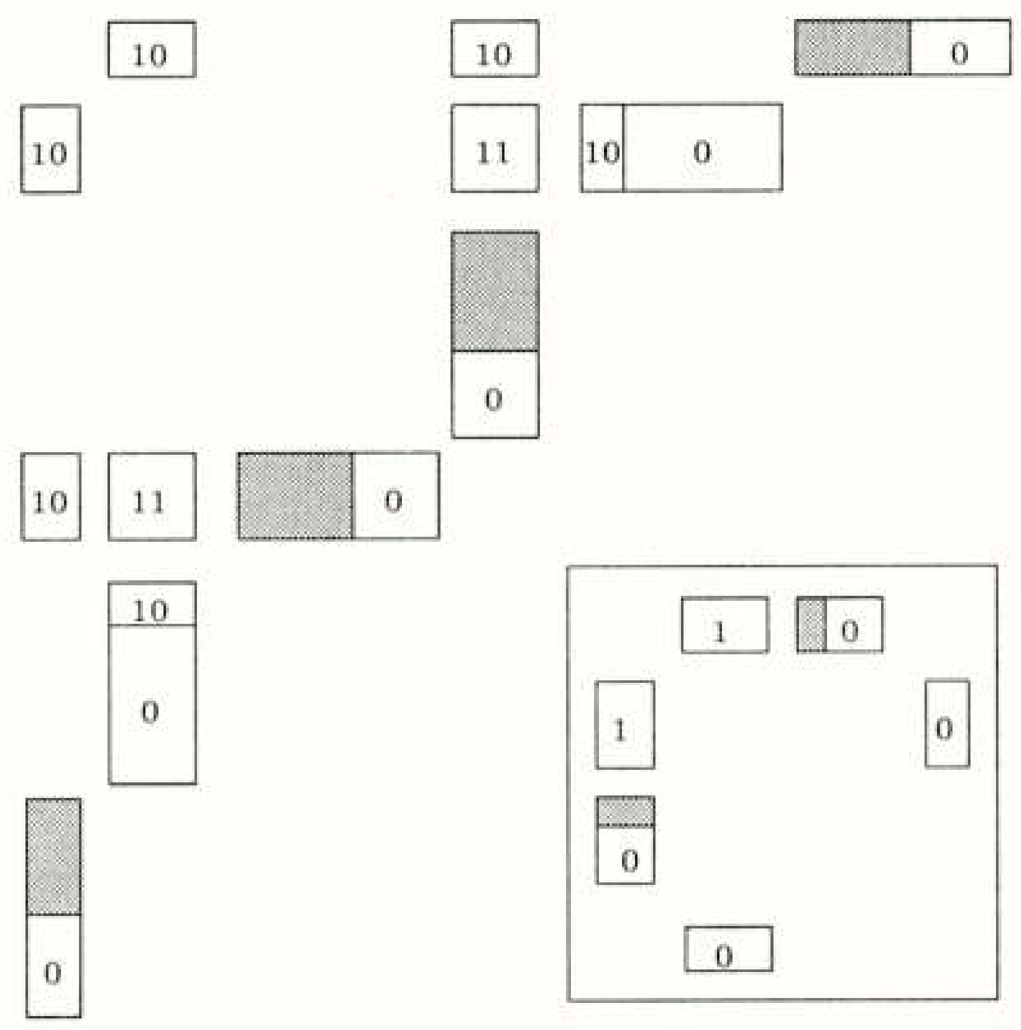


Fig. 5. The modification in the subdivision of a part of the outer rim.

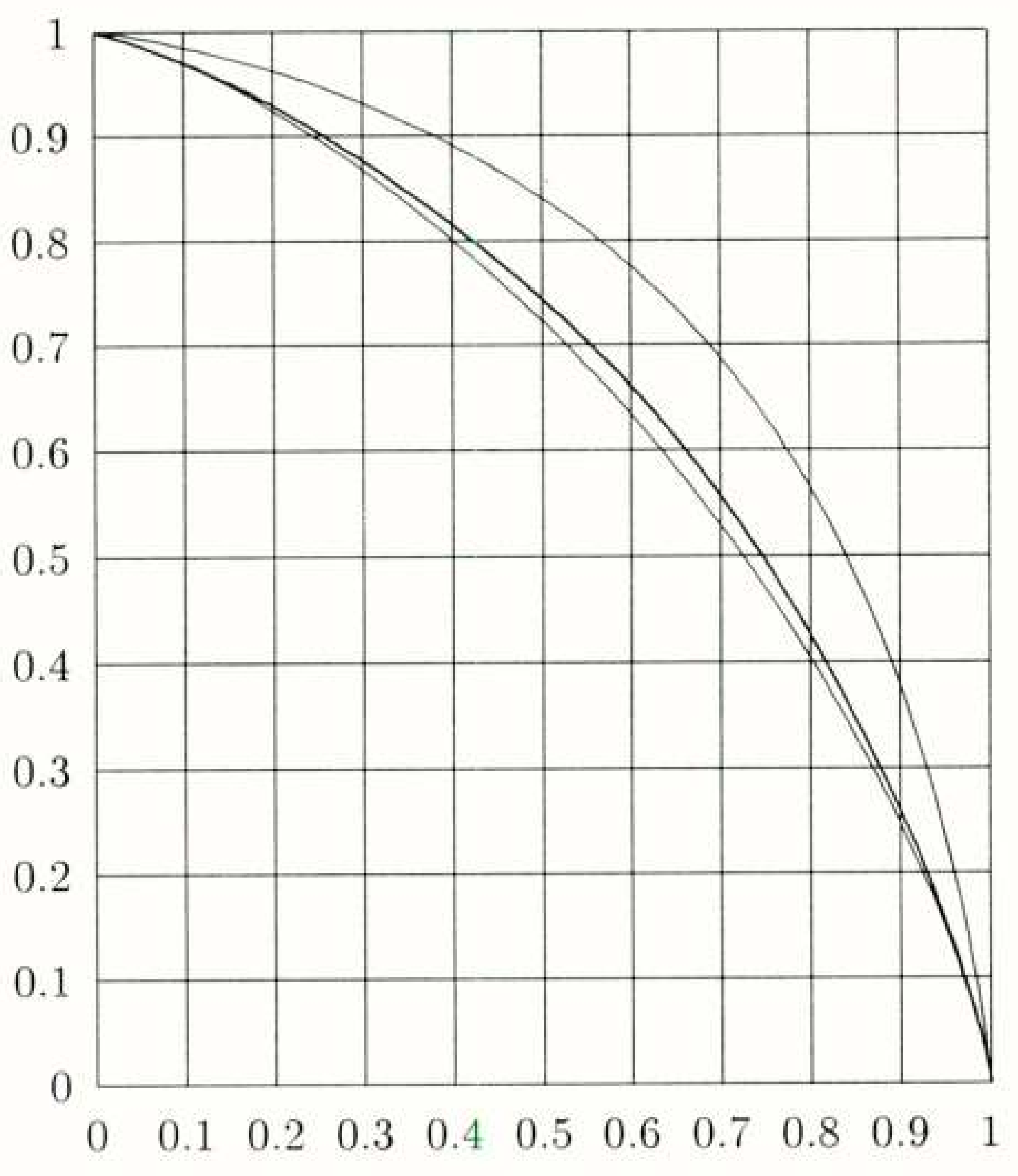


Fig. 6. The new achievable rate region.

C. The New Achievable Rate Region

The information rates in actual communication situations are not necessarily the same in both directions. The new coding strategy can also be adapted for the general unequal rates case. The inner and outer curves in Fig. 6 illustrate Shannon's inner and outer bound regions, respectively. The remaining curve shows the complete achievable rate region of the new coding strategy. Note that van Overveld's [12] results prove the operational achievability of the new rate region.

VI. Conclusions

This paper describes two new results. First, the class of discrete save-up strategies that achieve rates close to the best lower bound of the BMC, and that can be easily implemented is introduced. Second, a new achievable rate region for the BMC is established. The new coding strategy asymptotically achieves R_1=R_2=0.63072 bit per transmission in the case of symmetric operation. There are no reasons to conjecture optimality of the new strategy. However, a further substantial extension of the achievable rate region by unit square division is prohibitively difficult.

In general, information and communication theory is devoted to the fundamental limits of a communication system, and to techniques that approach these limits as close as possible. If the limits are unknown, or in other words, if the communication problem is not completely understood, then (i) it is impossible to express an opinion on the efficiency and complexity of the realization of a certain system, and (ii) it is possible that a certain solution is quickly out of date. To conclude, network information theory is still in its infancy, but if we consider the increasing importance of communication networks, then its relevance needs no further emphasis.

References

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. on Math. Statist. and Prob.*, vol. 1, 1961, pp. 611-644. Reprinted in *Key Papers in the Development of Information Theory*, D. Slepian, Ed. New York: IEEE Press, 1974, pp. 339-372.

[2] R. E. Blahut, *Digital Transmission of Information*. New York: Addison-Wesley, 1990.

[3] J. P. M. Schalkwijk, "The binary multiplying channel - a coding scheme that operates beyond the Shannon inner bound," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 107-110, Jan. 1982.

[4] J. P. M. Schalkwijk, "On an extension of an achievable rate region for the binary multiplying channel," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 445-448, May 1983.

[5] W. M. C. J. van Overveld, "Fixed-length strategies for the binary multiplying channel," *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 314-318, March 1988.

[6] J. P. M. Schalkwijk, "A new lower bound for the binary multiplying channel," in *Trans. 11th Prague Conference on Inform. Theory, Statistical Decision Functions, Random Processes*, 1992, pp. 341-347.

[7] Z. Zhang, T. Berger, and J. P. M. Schalkwijk, "New outer bounds to capacity regions of two-way channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 383-386, May 1986.

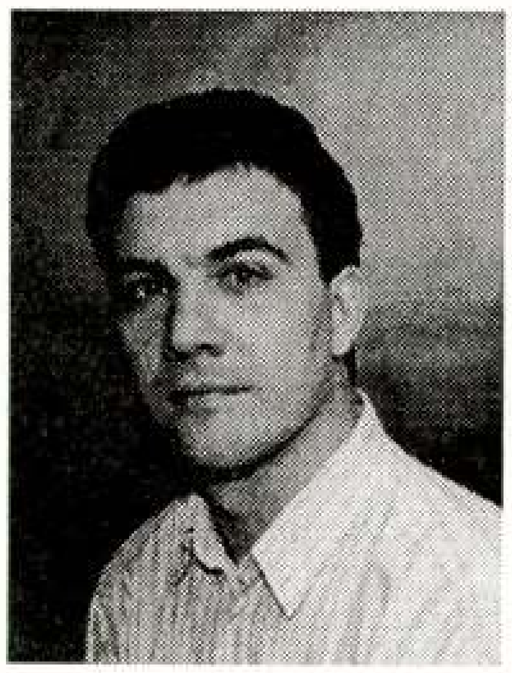
[8] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 44-53, Jan. 1989.

[9] J. P. M. Schalkwijk, H. B. Meeuwissen, and A. H. A. Bloemen, "Coding strategies and a new achievable rate region for the binary

multiplying channel," submitted to *IEEE Trans. Inform. Theory.*
[10] B. B. Mandelbrot, *The Fractal Geometry of Nature.* New York: Freeman and Company, 1983.
[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: John Wiley, 1991.
[12] W. M. C. J. van Overveld, *On the Capacity Region for Deterministic Two-Way Channels and Write-Unidirectional Memories.* Ph. D. dissertation, Department of Electrical Engineering, Eindhoven Univ. of Technol., Eindhoven, The Netherlands, 1991.
[13] J. P. M. Schalkwijk, "Some aspects of the information theoretic dialogue," *Discrete Math.*, vol. 106/107, pp. 407-413, 1992.

BASIS PRINCIPES IN (MPEG) VIDEOCODERING

Dr.ir. Reginald L. Lagendijk
Technische Universiteit Delft, Faculteit der Elektrotechniek,
Vakgroep Informatietheorie



Summary

The principles of (MPEG) video coding

The demand for economical high quality digital video television broadcasting necessitates an efficient usage of available bandwidth of the transmission channels. This justifies the worldwide efforts in techniques for digital video compression. In this presentation an overview will be given of the principles of video compression, namely spatial transforms (DPCM, DCT), motion compensated prediction, and quantization. Next, the basic configuration of a video compression system will be outlined, together with several of the options that exist to implement such a system. Finally, the MPEG-1 and MPEG-2 worldwide standardized compression methods will be discussed.

1. Introductie

In toenemende mate wordt visuele informatie in televisie en multimedia toepassingen in digitale vorm gerepresenteerd. Belangrijke voordelen hiervan zijn de robuustheid tegen transmissiefouten, de eenvoud tot manipulatie met digitale signaalprocessoren, en de synergie met computerapplicaties. De bandbreedte van digitale videosignalen is echter meestal veel groter dan de bandbreedte die feitelijk beschikbaar is voor de transmissie over digitale communicatielijnen (satellietlinks, ISDN, digitale informatiesnelweg). Een typerend voorbeeld is een videobronsignaal conform de CCIR-601 aanbeveling (zie Tabel I), dat een bandbreedte (of *bit rate*) heeft van 166 Mbs (megabit per seconde). Een dergelijk signaal zal getransporteerd moeten worden over kabelnetwerken, satelliet en aardse verbindingen met een bandbreedte van maximaal 10 Mbs. Om die reden is het noodzakelijk de bronsignalen te *comprimeren* of *coderen*, zodat voor de overdracht toch met de genoemde relatief kleine bandbreedte volstaan kan worden.

Het vakgebied van de *digitale videocodering* houdt zich bezig met het ontwikkelen van methoden en algoritmen die de compressie van videosignalen mogelijk maakt, veelal met factoren die liggen tussen de 5 voor eenvoudige methoden tot soms boven de 100 voor de meer complexe aanpakken [1-3]. Dit vakgebied staat reeds meer dan 20 jaar in de belangstelling van internationaal universitair onderzoek en, vooral door de opkomst van snelle digitale signaalverwerkingsapparatuur, ook reeds geruime tijd in die van de telecommunicatie-, computer- en consumentenindustrie. Vooral de laatste 5 jaar heeft door de opkomst van de pc, photo-cd, cd-rom, video-cd, cd-i, de digitale informatiesnelweg en digitale televisie een explosieve groei plaats gevonden van geïnteresseerde bedrijven en utilisanten van toepassingen waarin het gebruik van beeld- en videomateriaal essentiëel is.

In Tabel I zijn enkele representatieve digitale televisie bronformaten opgenomen, met daarbij de bandbreedte (bit rate in megabit per seconde), de beoogde bandbreedte van het digitale transmissiekanaal, en de hiermee samenhangende compressiefactor. We zien dat voor genoemde videoformaten de vereiste compressiefactor in de orde grootte van 10 tot 20 ligt. Dergelijke bandbreedte reducties kunnen niet bereikt worden door het toekennen van efficiënte variabele lengte codewoorden zoals Huffman codes [4], maar vereisen een aanpak waarbij compressie bereikt wordt ten koste van de kwaliteit van het gedecodeerde signaal. Met andere woorden, het ontvangen gecodeerde videosignaal is na decodering niet identiek aan het "oorspronkelijk" digitale videosignaal voorafgaand aan de bandbreedte-reducerende codering. We spreken in dit verband dan ook van niet-foutvrije codering.

De mate van verschil tussen het bronsignaal en het gedecodeerde signaal bepaalt de kwaliteit. Bij lage compressiefactoren is een hoge kwaliteit van het gedecodeerde signaal mogelijk, maar naarmate de compressiefactor opgevoerd wordt, zal de kwaliteit van het gedecodeerde signaal sterk teruglopen: compressiefactor en kwaliteit gaan altijd hand in hand. In Tabel I is daarom ook aangegeven wat grofweg de subjectieve kwaliteit is van de gecodeerde videosignalen ten opzichte van bestaande analoge opslag- en transmissieformaten.

In dit artikel wordt ingegaan op enkele van de meest gebruikte technieken voor compressie van digitale video. Dit zal leiden tot een basisschema dat de meeste gestandaardiseerde videocompressiesystemen (*video coders*) gemeenschappelijk hebben. Verschillende standaarden en zelfs verschillende implementaties van video coders binnen eenzelfde standaard kunnen echter sterk van elkaar verschillen. Er zal stil gestaan worden bij enkele van de belangrijkste opties in het ontwerpen van een video coder. Vervolgens wordt aandacht besteed aan de twee momenteel belangrijkste internationale videocompressiestandaarden, namelijk MPEG-1 en MPEG-2. Het artikel besluit met enkele speculaties over toekomstige ontwikkelingen op het gebied van de codering van digitale video.

2. Essentiële componenten in videocompressie

Bandbreedtereductie ofwel compressie van een willekeurig signaal is slechts mogelijk wanneer dit signaal redundantie bevat. Hiervoor kan bewijsvoering aangedragen worden uit de informatietheorie [4], maar in dit artikel wordt de voorkeur gegeven aan een minder strikte aanpak. We bespreken nu één voor één de belangrijkste componenten uit een videocompressiesysteem.

2.1 Representatie

Veel videoproduktieapparatuur zoals camera's en grafische ontwerppakketten representeren een (analoog) videosignaal in de RGB (rood-groen-blauw) kleurenruimte. Compressiemethoden werken vrijwel altijd met digitale videosignalen die gerepresenteerd worden in de YUV ofwel luminantie-chrominantie kleurenruimte. De luminantiecomponent Y is niets anders dan

format	*lines×pels per line (Y-UV comp.)*	*#pictures per sec*	*PCM band-width (Mbs)*	*channel band-width (Mbs)*	*compr. factor*	*quality level*
SIF	288×352 (144×176)	25 Hz (progr.)	30	1.15	25	Super VHS
CCIR-601 (tv signal)	576×720 (576×360)	50 Hz (2:1 interlace)	166	5-15	10 - 20	> Broadcast PAL
HD	1152×1920 (1152×960)	50 Hz (progr.)	$>10^3$	80	10 - 20	

Tabel I: Enkele digitale videobronformaten en bandbreedten, beoogde transmissiebandbreedten, vereiste compressiefactoren, en de hiermee gepaard gaande subjectieve kwaliteiten.

de zwart-wit informatie in een videosignaal, terwijl U en V signalen de kleurverschilsignalen R-Y en B-Y zijn. Luminantie en chrominantiesignalen kunnen via een matrixoperatie uit de RGB signalen verkregen worden. Het blijkt nu dat U en V signalen aanzienlijk minder energie en hoogfrequent informatie bevatten dan het Y signaal, en dat U en V signalen van minder belang zijn bij de menselijke kwaliteitsbeoordeling van videosignalen. Om die reden worden chrominantiesignalen gerepresenteerd met een kleinere horizontale en vaak ook vertikale bandbreedte. Dit wordt bereikt door na de RGB-YUV transformatie de U en V signalen horizontaal en vertikaal banddoorlaat te filteren en onder te bemonsteren. Tabel II toont voor een videosignaal met CCIR-601 resolutie het aantal lijnen en aantal beeldpunten per lijn in de RGB representatie, in de zogenaamde 4:2:2 YUV representatie en in de 4:2:0 YUV representatie. Deze laatste representatievorm wordt vrijwel altijd gebruikt in video compressiemethoden. Merk op dat ten opzichte van RGB, de 4:2:0 YUV representatie reeds 50% bandbreedtereductie levert. Elk van de componenten wordt aanvankelijk in 8 bit per bemonstering gerepresenteerd door middel van PCM codering (puls-code-modulatie) [1]

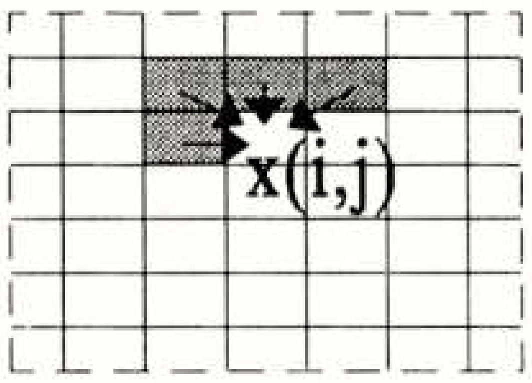
representatie	*component*	*#lijnen*	*#beeldpunten per lijn*	*#beeld punten per frame* (10^6)
RGB	R, G, B	576	720	1.24
4:2:2 YUV	Y	576	720	
	U, V	576	360	0.83
4:2:0 YUV	Y	576	720	
	U, V	288	360	0.62

Tabel II: Afmetingen (in lijnen en beeldpunten per lijn) van een RGB, 4:2:2 YUV en 4:2:0 YUV digitaal videosignaal.

2.2 *Intraframe codering*

Beelden of *frames* bevatten veel gestructureerde en dus voorspelbare spatiële (d.w.z. binnen een beeld) gebieden: juist aan deze structuren koppelt de mens zijn interpretatie van de beeldinformatie. Vanuit compressie oogpunt zijn echter voorspelbare structuren redundant, dat wil zeggen dat alles dat voorspelbaar is, niet getransporteerd zou hoeven te worden van zender naar ontvanger, waarmee bandbreedte reductie mogelijk wordt.

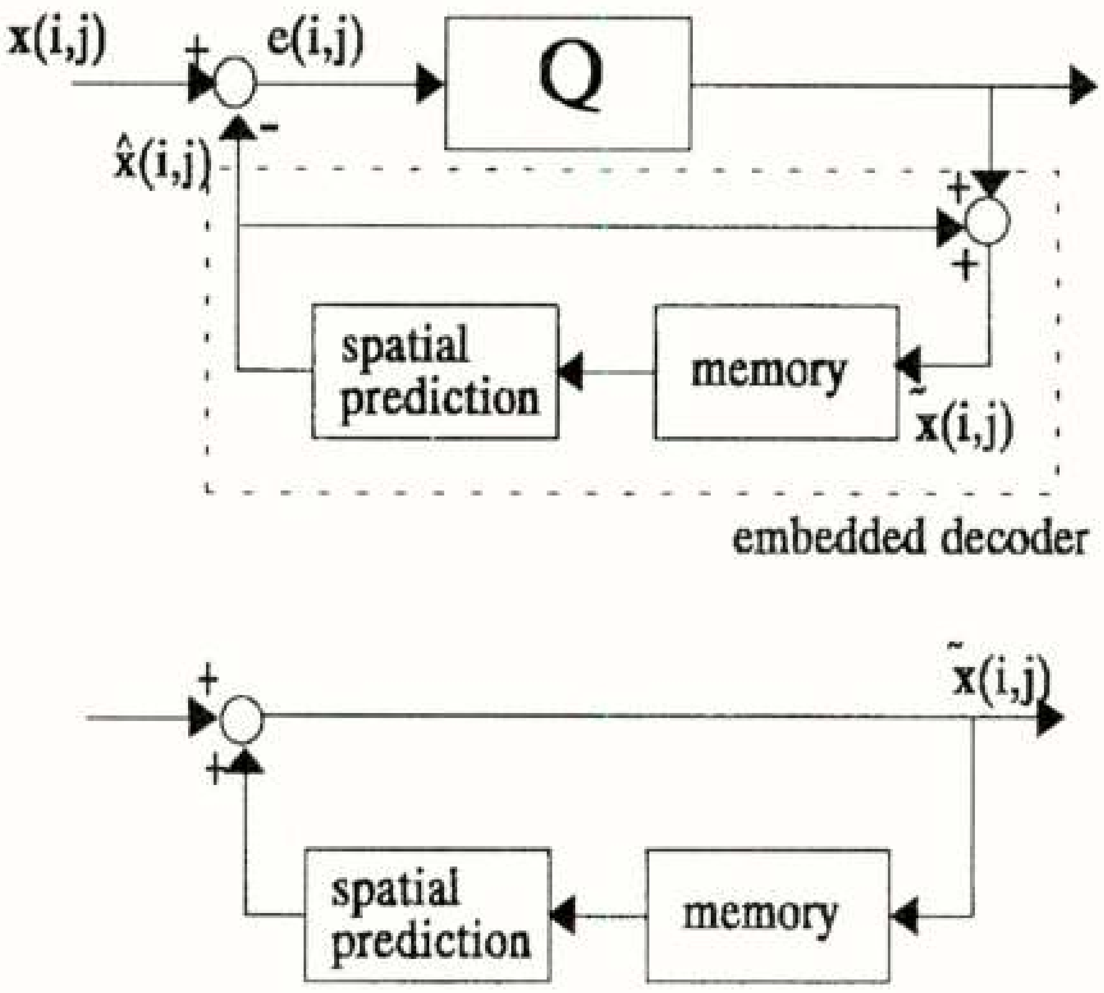
In de digitale videocodering wordt vaak gebruik gemaakt van twee mechanismen om de spatiële redundanties aan de encoderzijde te verwijderen uit frames, namelijk spatiële *differentiële* PCM (*DPCM*) en de *discrete cosinus transformatie* (*DCT*). We gaan nu kort in op beide methoden, die aangetroffen worden in vrijwel alle gestandaardiseerde beeld- of videocoderingsmethoden. Coderingsmethoden die de spatiële redundantie uit een beeld verwijderen worden aangeduid met intraframe coders.



Figuur 1: Spatiële predictie in een DPCM systeem.

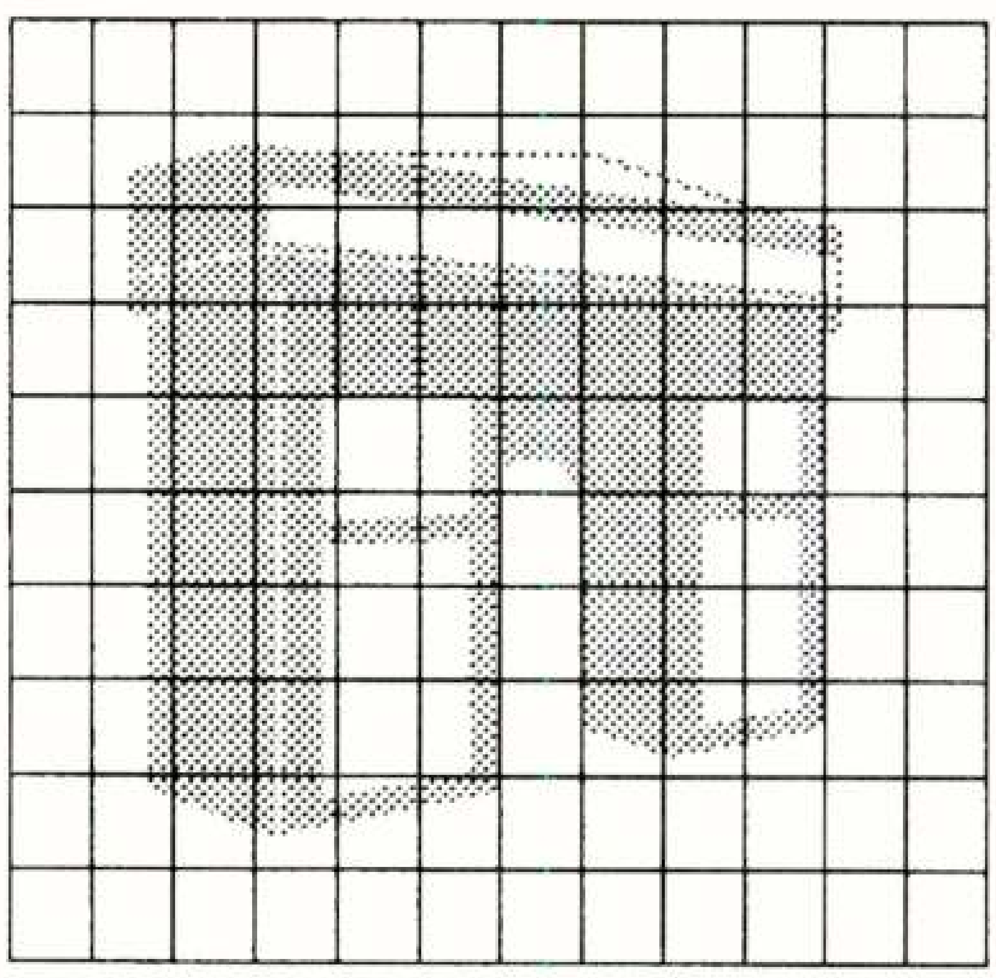
Spatiële DPCM veronderstelt dat alle beeldpunten uit een frame via het raster sequentiëel geordend zijn. Dat wil zeggen dat het frame van linksboven naar rechtsonder afgetast wordt. Stel nu dat in de coder reeds de signaalwaarden (helderheid, luminantie, chrominantie) van de beeldpunten in het grijze gebied in Figuur 1 gecodeerd zijn. Op grond van deze gecodeerde, en dus getransporteerde en bij de decoder bekende informatie, kan een voorspelling worden gemaakt van de signaalwaarde van het beeldpunt $x(i,j)$. Deze voorspelling wordt opgebouwd uit een gewogen som van reeds gecodeerde beeldpunten. Uiteraard zal de voorspelling niet precies gelijk zijn aan de fei-

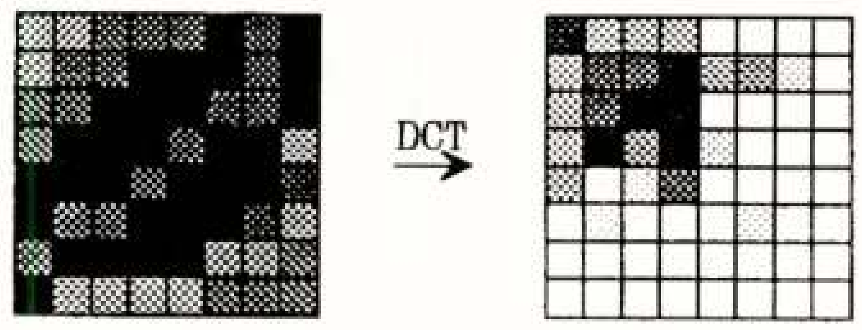
telijk signaalwaarde van het beeldpunt: de grootte van de voorspellingsfout $e(i,j)$ zal afhangen van de nauwkeurigheid van de voorspelling. Wel is het nu zo dat als de voorspelling op de juiste manier gebeurt, de voorspellingsfout niet langer redundant is: slechts deze informatie wordt dan ook, na kwantisatie (Q), naar de decoder getransporteerd via het kanaal. Figuur 2 laat de structuur van de resulterende DPCM encoder en decoder zien. Hierbij valt het op dat in de encoder in feite een decoder opgenomen is. Om deze reden wordt veelal slechts alleen de encoder getoond. De compressie- versus kwaliteitsprestatie van een spatiëel DPCM systeem is vrij beperkt, en hangt af van hoe de spatiële voorspelling wordt gemaakt en van de eigenschappen van de kwantisator Q.



Figuur 2: (a) Intraframe encoder en (b) decoder gebaseerd op spatiële DPCM

Een andere wijze om de spatiële redundantie te kunnen verwijderen is via een lineaire *decorrelerende* transformatie van de beeldinhoud. Decorreleren wil niets anders zeggen dan dat de voorspelbaarheid tussen een groepje beeldpunten onderling verminderd of zelfs tot nul teruggebracht wordt. Uit complexiteit oogpunt wordt een dergelijke transformatie uitgevoerd op kleine subbeelden, bijvoorbeeld 8x8 of 16x16 blokjes. De optimaal decorrelerende transformatie voor dergelijke subbeelden is de z.g. Karhunen-Loeve transformatie, die voor praktische toepassingen echter veel te ingewikkeld is. Een

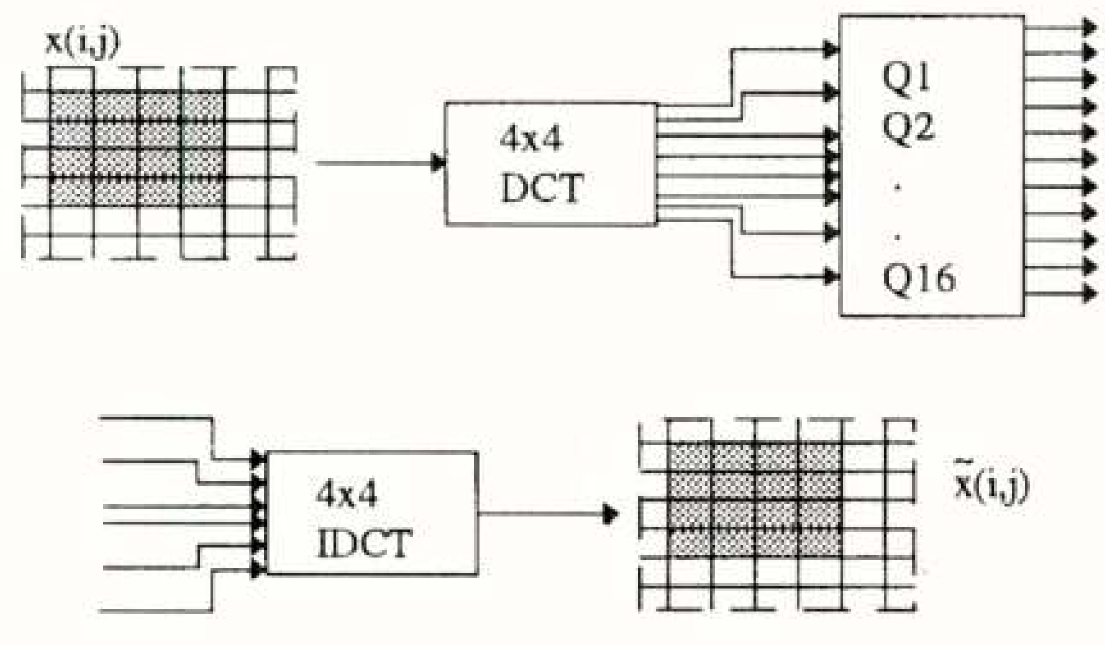




Figuur 3: Illustratie DCT transformatie.

goede benaderende decorrelerende transformatie is de discrete cosinus transformatie (DCT), wat in feite een vereenvoudige (discrete) Fourier transformatie (DFT) is waarbij alleen cosinus termen worden toegelaten.

Eerst wordt nu een frame opgebroken in aansluitende 8x8 subbeelden. Vervolgens worden de signaalwaarden in elk subbeeld gedecorreleerd door middel van het toepassen van een twee-dimensionale DCT. Dit resulteert, voor ieder subbeeld, in 8x8 DCT *coefficiënten* (zie Figuur 3) die veel overeenkomsten vertonen met DFT *spatiële* frequentiecomponenten. Bijvoorbeeld, één van de DCT coefficiënten is gelijk aan de gemiddelde waarde in een subbeeld. Omdat een beeld voornamelijk laagfrequent spatiële informatie bevat, zullen "laag-frequent" DCT coefficiënten aanzienlijk grotere waarden hebben dan "hoog-frequent" DCT coefficiënten. Tevens zijn voor de visuele kwaliteit lagere spatiële frequenties van veel groter belang dan hogere frequenties. Om deze twee redenen worden "laag-frequent" DCT coefficiënten, en vooral de gemiddelde waarde, met veel grotere nauwkeurigheid naar de decoder getransporteerd dan "hoog-frequent" DCT coefficiënten. Dit wordt bereikt door verschillende kwantisatoren te gebruiken voor de verschillende DCT coefficiënten (zie Figuur 4). Effectief betekent dit dat "laag-frequent" DCT coefficiënten met een hogere bit rate gecodeerd worden dan "hoog-frequent" DCT coefficiënten. Sommige DCT coefficiënten worden zelfs simpelweg op 0 afgerond, en hoeven dan is het geheel niet naar de decoder gestuurd te worden: dit levert een aanzienlijk compressie op.



Figuur 4: (a) Intraframe encoder en (b) decoder gebaseerd op een (4x4) DCT transformatie.

Aan de decoderzijde worden de ontvangen gekwantiseerde DCT coefficiënten teruggeplaatst in een 8x8 subbeeld, waarop vervolgens de inverse DCT uitgevoerd wordt. Doordat de DCT een orthogonale transformatie is, is de inverse DCT operatie vrijwel identiek aan de DCT zelf, wat aantrekkelijk is vanuit implementatie oogpunt. Alhoewel de DCT is een efficiënte transformatiemethode is, hangt de uiteindelijk prestaties van een compressiesysteem gebaseerd op de DCT toch vooral af van de keuze van de kwantisatoren voor de verschillende DCT coefficiënten, en hun in samenhang gekozen onderlinge bit rates en dus nauwkeurigheden.

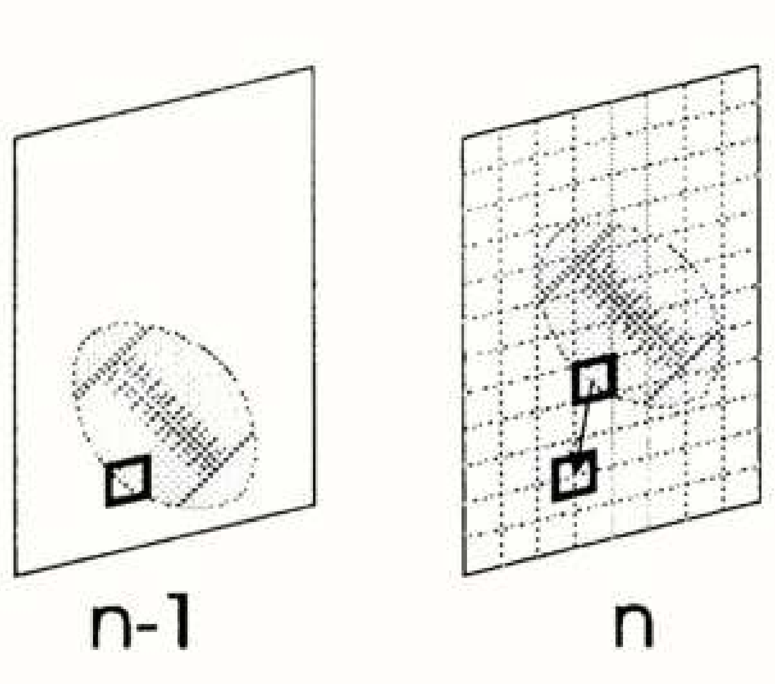
Alhoewel de DCT wel de onderlinge afhankelijkheid binnen een subbeeld verwijdert, wordt daarmee niet de voorspelbaarheid tussen naburige subbeelden verminderd. Om die reden wordt soms op de individuele DCT coefficiënten nog DPCM toegepast, waarblj de voorspelling van een DCT coefficiënt wordt opgebouwd uit corresponderende DCT coefficiënten uit naburige subbeelden. Een dergelijke differentiële codering levert vooral een aanzienlijke compressie op voor de DCT coefficiënt die de gemiddelde grijswaarde in een subbeeld representeert.

2.3 *Interframe codering*

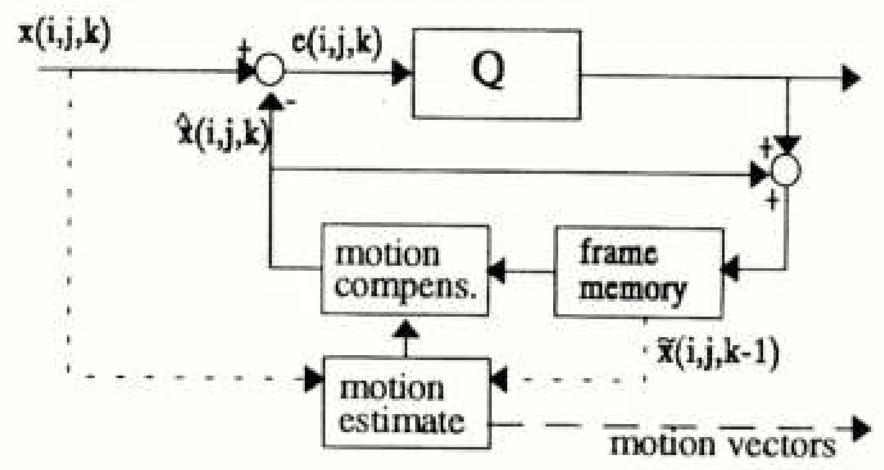
Videosignalen zijn niet alleen redundant in spatiële zin, maar ook in de tijd. Op grond van een enkel videoframe is het vrijwel altijd mogelijk te voorspellen hoe het volgende videoframe er uit zal zien. Om deze redundantie in de tijd te verwijderen, wordt *temporele DPCM* toegepast. Hierbij wordt aan de hand van een reeds gecodeerd frame het volgende frame voorspeld, en wordt slechts de voorspellingsfout naar de decoder getransporteerd. Als er weinig verandering is tussen opeenvolgende frames, is het voldoende om als voorspelling voor het volgende frame het huidige reeds gecodeerde frame te nemen. Echter, als zich veranderingen ten gevolge van bewegende objecten of camerabewegingen voordoen, moet hiermee rekening gehouden worden in het maken van de temporele voorspelling. Hiertoe is wel eerst de bepaling van de beweging in het videosignaal noodzakelijk. Dit probleem staat bekend als *bewegingsschatting*.

Bewegingsschatting is een onderwerp dat in meer applicatiegebieden voorkomt, zoals bijvoorbeeld bij bewegingsanalyse voor medische en industriële toepassingen. Er is dan ook een veelheid aan technieken voor bewegingsschatting beschikbaar, en regelmatig worden nieuwe methoden voorgesteld. Uit complexiteitsoogpunt wordt binnen de videocompressie meestal gebruik gemaakt van methoden die de verplaatsing schatten van een 8x8 of 16x16 subbeeld ten opzichte van een reeds gecodeerd vorig frame. Figuur 5 illustreert hoe in het beeld n-1 een corresponderend subbeeld wordt gevonden voor het te coderen beeld n. Het verschil in positie tussen de twee subbeelden noemt men de *bewegings-* of *verplaatsingsvector*, die (veelal in gecodeerde vorm) naar de decoder gezonden dient te worden om een juiste decodering mogelijk te maken. Na het bepalen van de bewegingsvector voor alle subbeelden wordt een voor beweging gecompenseerde voorspelling gemaakt, die vervolgens afgetrokken wordt van het te coderen beeld. Slechts dit niet-redundante bewegingsgecompenseerde verschilsignaal wordt nu gekwantiseerd en naar de decoder verzonden. Het resulterende *interframe* coderingsschema wordt getoond in Figuur 6.

Bewegingsgecompenseerde temporele predictie wordt in de meeste videocoderingssystemen toegepast. De mate van compressie die hiermee bereikt kan worden is vooral afhankelijk van de toegepaste bewegingsschatter.



Figuur 5: Illustratie van bewegingsschatting

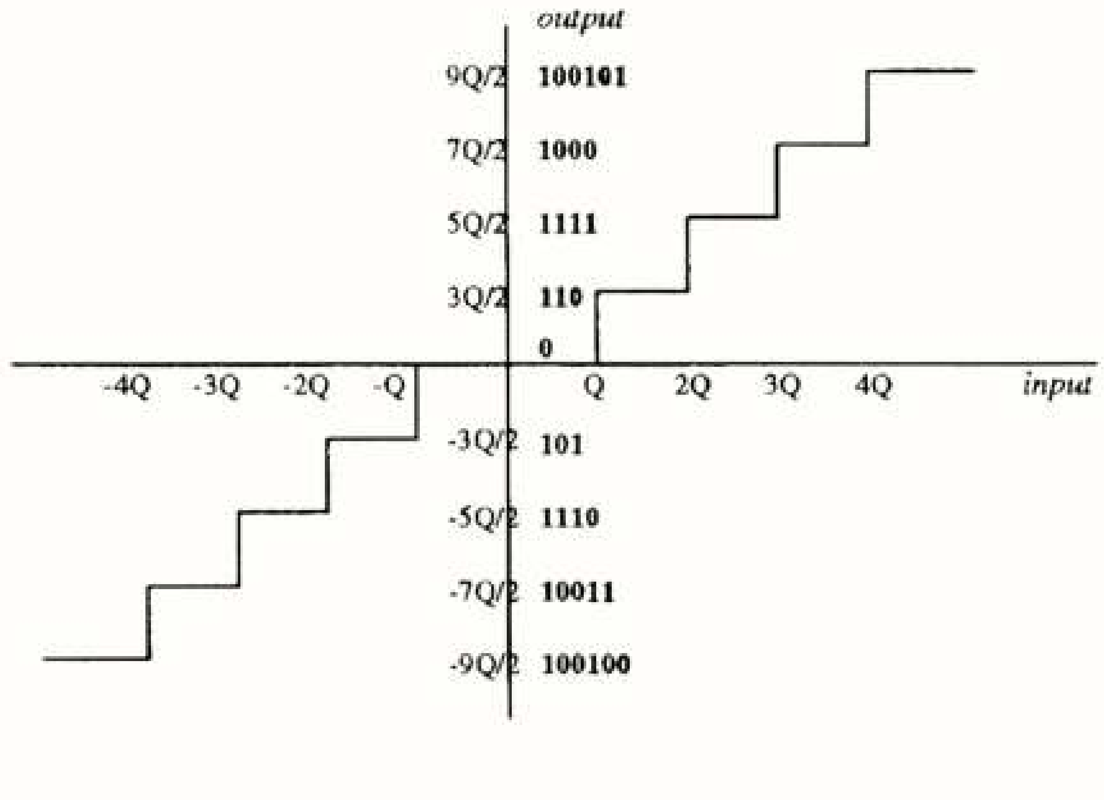


Figuur 6: Interframe coderingsschema gebaseerd op temporele DPCM.

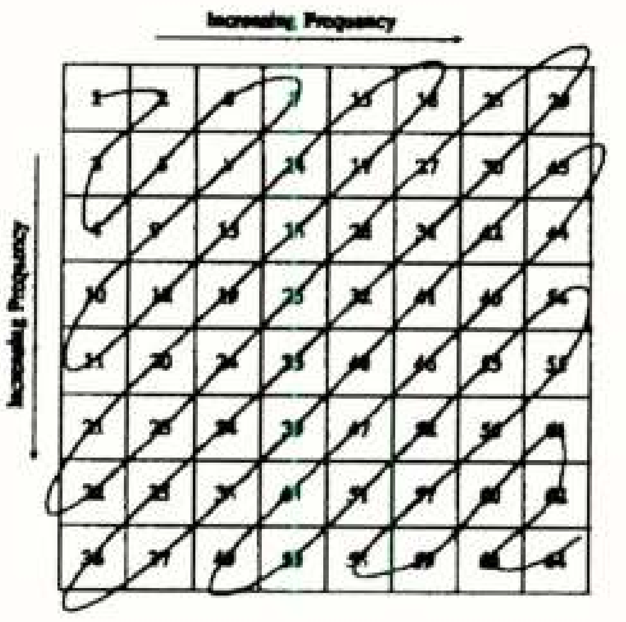
2.4 *Kwantisatie en variabele lengte codering*

Elk compressiesysteem zal een kwantisator moeten bevatten om feitelijke reductie van de vereiste transmissiebandbreedte te bewerkstelligen. Zonder kwantisator zijn de schema's in Figuur 2, 4 en 6 slechts lineaire bewerkingen op de videosignalen, die geen enkele bandbreedtereductie teweeg brengen. Feitelijk zijn DPCM, de DCT en bewegingsgecompenseerde temporele predictie slechts methoden om een geschikte niet-redundante representatie te vinden voor het videosignaal. Deze niet-redundante representatie dient dan in nog wel in gekwantiseerde vorm naar de decoder getransporteerd te worden.

Een veel gebruikte kwantisator is de uniforme kwantisator met dode-zone. Figuur 7 laat een voorbeeld zien van een dergelijke kwantisator. De ingangswaarden (bijvoorbeeld DCT coefficiënten) worden afgebeeld via de kwantisatorcurve op een eindig aantal discrete uitgangswaarden (vertikale as), welke vervolgens worden gerepresenteerd door codewoorden met variabele lengte (VLC). Deze codewoorden worden ontworpen met behulp van bijvoorbeeld het Huffman coderingsrecept. We zien dat veel voorkomende kleine signaalwaarden gerepresenteerd worden met relatief korte codewoorden: dit resulteert in een lage bit rate. Grotere signaalwaarden, die relatief weinig voorkomen, zullen gerepresenteerd worden met langere codewoorden. Het aantal bits dat voor een gegeven signaal uiteindelijk geproduceerd wordt door de kwantisator hangt af van de instelling van de schaalparameter Q. Heeft Q een grote waarde, dan worden veel signaalwaarden op korte codewoorden afgebeeld, wat een lage bitproduktie tot gevolg heeft. Tevens zal dan het gemiddelde verschil tussen ingangswaarden en gecodeerde waarden groot zijn, d.w.z. een lage bit rate gaat gepaard met een grote gemiddelde kwantisatiefout. Aan de andere kant, een kleine waarde van Q levert juist een grote bitproduktie op en een kleine gemiddelde kwantisatiefout.



figuur 7a



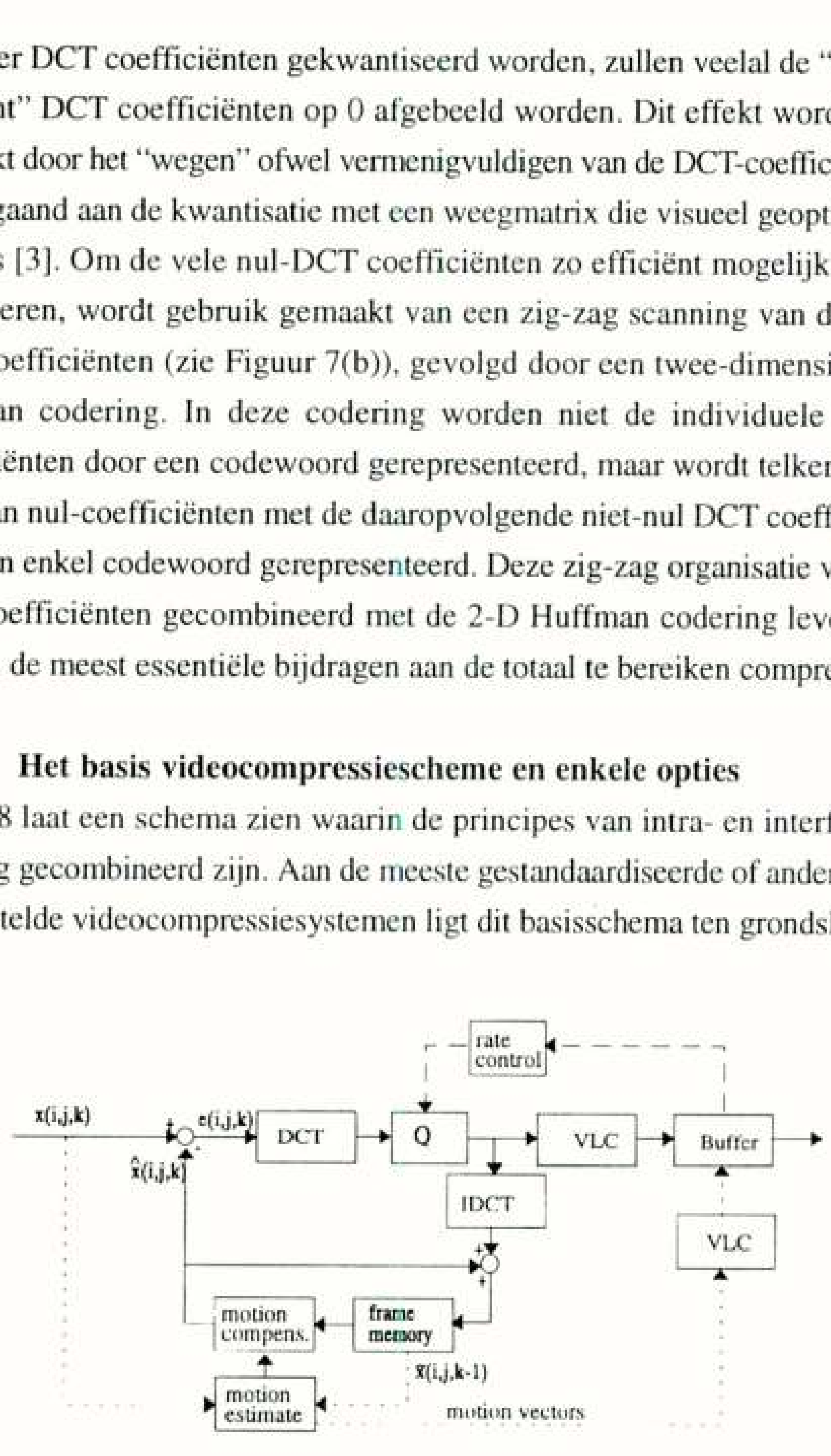
figuur 7b

Figuur 7: (a) Uniforme kwantisator met dode zone en schaalparameter Q, *(b) zig-zag scanning van gekwantiseerde DCT coefficiënten.*

Wanneer DCT coefficiënten gekwantiseerd worden, zullen veelal de "hoogfrequent" DCT coefficiënten op 0 afgebeeld worden. Dit effekt wordt nog versterkt door het "wegen" ofwel vermenigvuldigen van de DCT-coefficiënten voorafgaand aan de kwantisatie met een weegmatrix die visueel geoptimaliseerd is [3]. Om de vele nul-DCT coefficiënten zo efficiënt mogelijk te representeren, wordt gebruik gemaakt van een zig-zag scanning van de 8x8 DCT coefficiënten (zie Figuur 7(b)), gevolgd door een twee-dimensionale Huffman codering. In deze codering worden niet de individuele DCT coefficiënten door een codewoord gerepresenteerd, maar wordt telkens een serie van nul-coefficiënten met de daaropvolgende niet-nul DCT coefficiënt door een enkel codewoord gerepresenteerd. Deze zig-zag organisatie van de DCT coefficiënten gecombineerd met de 2-D Huffman codering levert de één van de meest essentiële bijdragen aan de totaal te bereiken compressie.

3. Het basis videocompressiescheme en enkele opties

Figuur 8 laat een schema zien waarin de principes van intra- en interframe codering gecombineerd zijn. Aan de meeste gestandaardiseerde of anderszins voorgestelde videocompressiesystemen ligt dit basisschema ten grondslag.



Figuur 8: Basisschema waarin intra- en interframe codering gecombineerd worden.

Het coderen van frame $x(i,j,k)$ gaat als volgt in zijn werk. Op grond van het vorige gecodeerde frame $x(i,j,k-1)$ wordt een bewegingsgecompenseerde voorspelling gemaakt. De voorspellingsfout $e(i,j,k)$ is temporeel niet langer redundant, maar nog wel spatiëel. Daarom wordt $e(i,j,k)$ DCT getransformeerd en worden de DCT coefficiënten gekwantiseerd en zig-zag gescanned om uiteindelijk gerepresenteerd te worden door codewoorden met variabele lengte. De encoder berekent zelf ook een gedecodeerde versie van frame $x(i,j,k)$ zodat dit resultaat door encoder en decoder gebruikt kan worden bij de voorspelling van het volgende frame $x(i,j,k+1)$.

Bovenstaand basisschema kan worden uitgebreid met talloze opties en toevoegingen die de uiteindelijke compressie-kwaliteit verhouding zullen bepalen. We bespreken hier een drietal belangrijke opties.

3.1 Bit rate regeling

He schema in Figuur 8 produceert over het algemeen meerdere bitstromen, bijvoorbeeld de VLC gecodeerde DCT coefficiënten, de (gecodeerde) bewegingsvectoren, en meestal ook controle- en synchronisatieinformatie. Deze stromen worden in een *uitgangsbuffer* gemultiplexed tot de uiteindelijk *video bitstroom* ofwel *elementary bitstream*. Zonder verdere voorzorgsmaatregelen zal de hoeveelheid bits die per seconde geproduceerd wordt, sterk kunnen variëren. Dit hangt bijvoorbeeld af van de temporele en spatiële voorspelbaarheid van het te coderen video materiaal. De meeste transmissiekanalen vragen echter om een constante bit rate. Hiertoe is het noodzakelijk de videocoder zo te regelen dat aan de uitgang van de buffer ten alle tijden een constante bit stroom uitgegeven kan worden: overlopen of leeglopen van de buffer moet worden voorkomen.

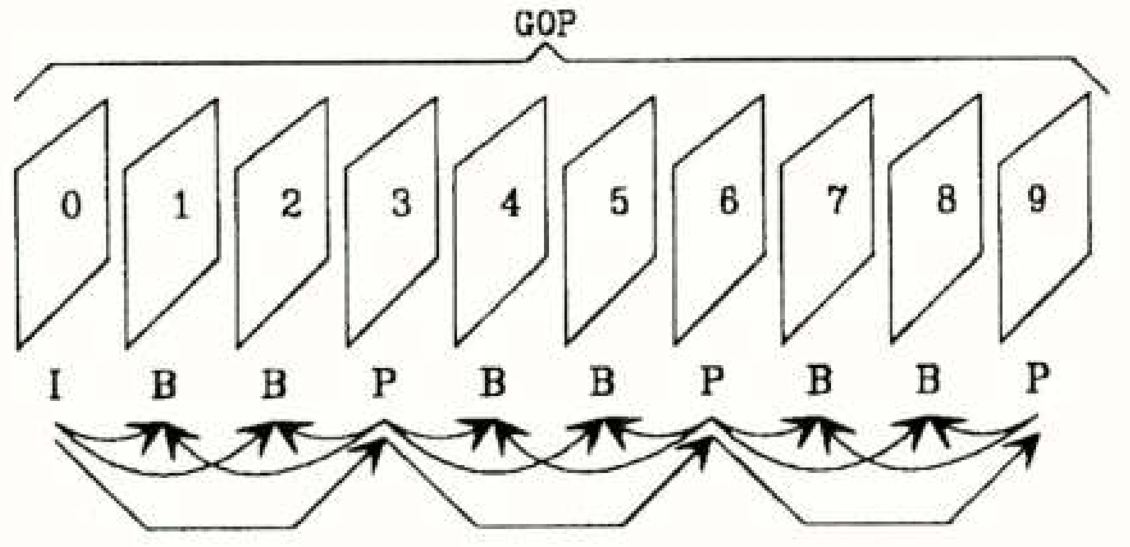
Regeling van de coder en de geproduceerde bit rate kan eigenlijk alleen gebeuren door de schaalparameter Q van de kwantisator te regelen. Als de uitgangsbuffer vol dreigt te lopen, wordt Q vergroot zodat er minder bits per seconde wordt geproduceerd, terwijl als de buffer leeg dreigt te lopen Q verkleind kan worden. Een dergelijke achterwaartse (teruggekoppelde) regeling van de kwantisator heeft als nadeel dat ook de kwaliteit van het gecodeerde beeld drastisch kan variëren van frame tot frame en zelfs binnen een frame. Enige vooruitregeling, waarbij een geschikte waarde van de schaalparameter Q voorspeld wordt vanuit het verleden, is hierbij dan ook noodzakelijk. Meer complexe bit rate regelingen maken zelfs gebruik van beeld-analyse methoden en visuele criteria om een geschikte waarde van Q voorafgaand aan de kwantisatie te kunnen bepalen.

3.2 Predictiemodes

In het schema in Figuur 8 wordt elk frame op grond van het direkt daaraan voorafgaande frame voorspeld. In een praktische coder is predictie niet altijd gewenst, of zijn andere vormen van predictie soms meer efficiënt. Dit argument heeft uiteindelijk geleid tot de introductie van 3 predictiemodes in de meeste compressiesystemen, namelijk:

- *Intra-coded (I) frames*: deze frames worden gecodeerd zonder temporele predictie, en kunnen dus onafhankelijk van andere frames gedecodeerd worden. Een op deze wijze gecodeerd frame kan als *entrypoint* in een gecodeerde video bitstroom gebruikt worden, wat bijvoorbeeld noodzakelijk is bij het *editten* van gecodeerde video maar ook bij het opstarten van een decoder bij het wisselen van televisiekanalen. Tevens is het van uit compressie oogpunt gezien gewenst om een I-frame te positioneren direkt na elke scene wisseling.
- *Predicted (P) frames*: deze frames worden temporeel voorspeld vanuit een voorafgaand frame, precies zoals aangegeven is in Figuur 8. P-frames vergen uiteraard een lager aantal bits dan I-frames.
- *Bidirectionally predicted (B) frames:* hierbij wordt een frame voorspeld vanuit een temporeel voorgaand *en* temporeel volgend frame. In de tijd gezien wordt een B-frame dus vanaf twee zijden (bi-directioneel) voorspeld. Het voordeel van een dergelijke predictie is dat occluderende beeld informatie beter voorspeld kan worden, en daarmee de bit rate voor B-frames aanzienlijk lager ligt dan die voor I en P-frames. Het nadeel is wel dat er meer bewerkingen nodig zijn, en dat de volgorde van transmissie van de frames niet langer hetzelfde is als de oorspronkelijke tijdsvolgorde, wat extra buffering vergt in encoder en decoder.

Figuur 9 laat een voorbeeld zijn van hoe I, P en B frames georganiseerd kunnen worden in een zich in de tijd herhalende *Group of Pictures* (GOP) structuur. In dit voorbeeld is de feitelijke codering, transmissie en decoderingsvolgorde als volgt: frame 0 (I), 3 (P), 1 (B), 2 (B), 6 (P), 4 (B), 5 (B), 9 (P), 7 (B), 8 (B), gevolgd door het I frame van de volgende GOP.



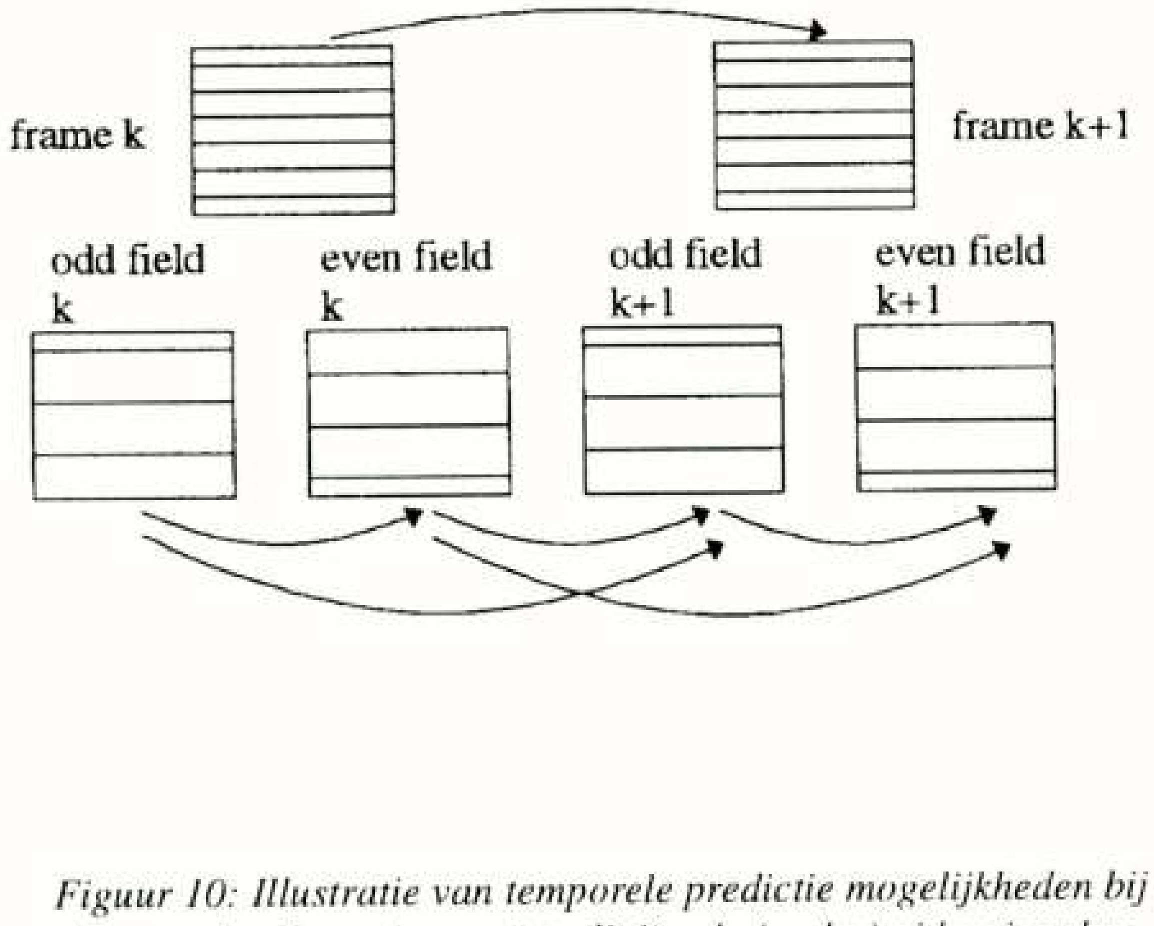
Figuur 9: Voorbeeld van een GOP structuur.

Belangrijk is hierbij op te merken dat de meeste coders *fallback* mogelijkheden hebben in het geval een predictiemode lokaal niet voldoende goed werkt. Bijvoorbeeld, als in een P-frame plaatselijk geen juiste temporele predictie mogelijk blijkt te zijn, dan kan daar de beeldinformatie in I-mode (zonder referentie naar andere frames) gecodeerd worden.

3.3 Frame- versus field-gebaseerde predictie

Wanneer we te maken hebben met geïnterliniëerde video volgens de CCIR601 norm, bestaat een enkel frame uit twee fields: het eerste field bevat de oneven lijnen uit het beeld, het tweede field de even lijnen. Het is hierbij van belang te realiseren dat de twee fields van verschillende tijdstippen afkomstig zijn, die 20 msec uit elkaar liggen. Dit heeft twee gevolgen voor een coderingssysteem. In de eerste plaats zal ten gevolge van horizontale beweging in de scene een verschuiving van beeldinformatie zichtbaar zijn tussen de even en oneven lijnen. Dit introduceert kunstmatige vertikale frequenties in een frame, wat zal leiden tot grote vertikale DCT coefficiënten wat weer tot verminderde compressie-efficiëntie leidt. Het lijkt voor de hand te liggen om de coder dan niet op frames te baseren, maar op fields. Dit vermindert echter weer de coderingsefficiëntie in gebieden waar geen horizontale beweging optreedt.

Een tweede gevolg van interliniering is dat bij de temporele voorspelling nu meerdere mogelijkheden ontstaan. Waar een geheel frame slechts voorspeld kan worden op grond van een geheel voorgaand frame, kan bijvoorbeeld een *even* field voorspeld worden op grond van een vorig *even* field of op grond van een vorig *oneven* field (Figuur 10). Het is niet mogelijk hierbij vooraf een keuze te maken. Met andere woorden, coders die met geïnterliniëerde videosignalen werken moeten adaptief kunnen bepalen welk van de temporele predicties het meest efficiënt is.



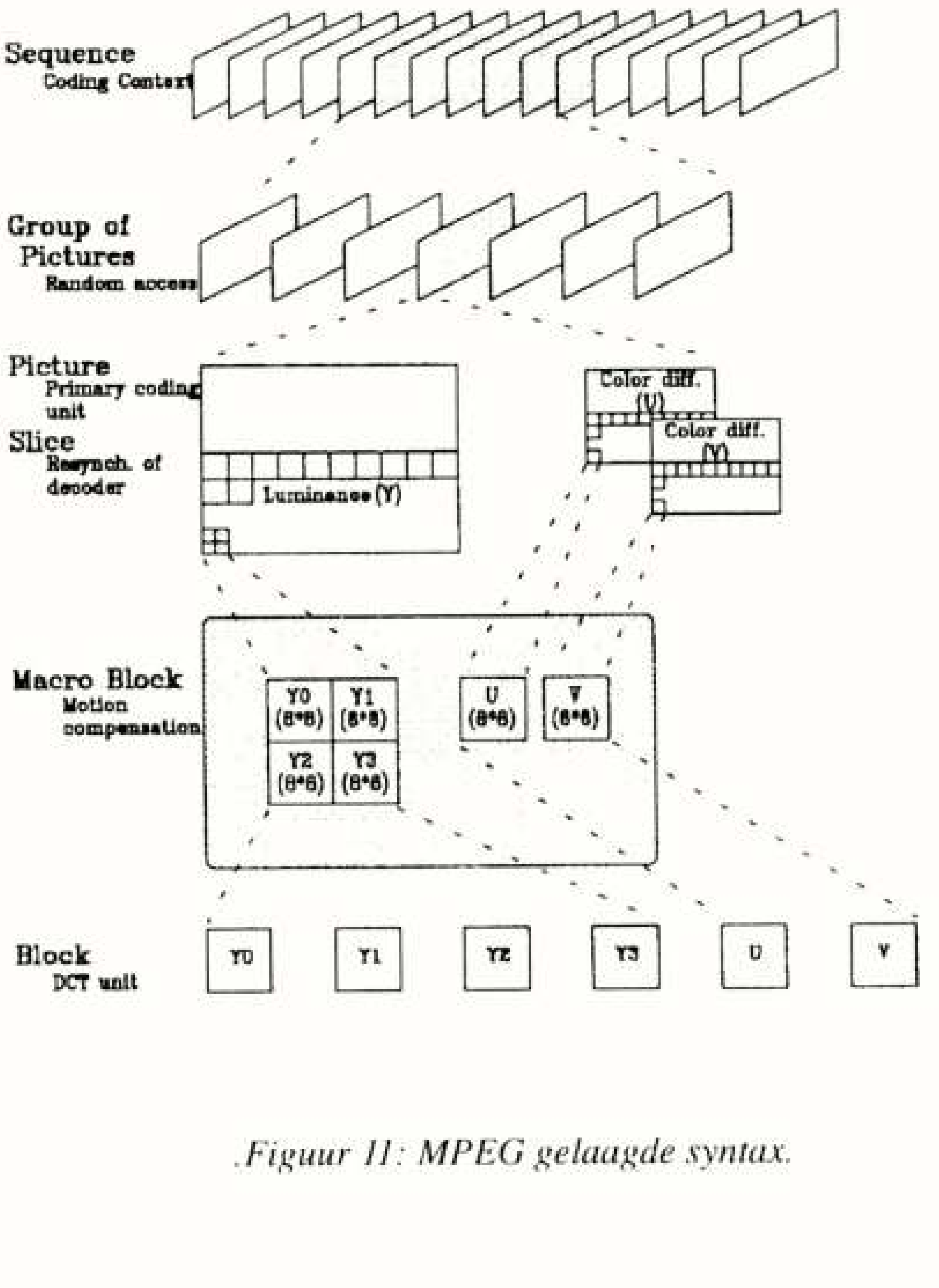
Figuur 10: Illustratie van temporele predictie mogelijkheden bij progressive (boven) en geïnterliniëerde (onder) videosignalen.

4. MPEG video coderingssystemen

De snelle opkomst van videocoderingstechnieken en de wens tot een wereldwijde standaard voor codering van videosignalen (in eerste instantie ten behoeve van digitale opslag media, later ook voor transmissietoepassingen), heeft geleid tot de MPEG standaardisatie. MPEG (Moving Pictures Experts Group) is het acroniem voor een ISO/IEC werkgroep (JTC1 SC2 WG11) waarin een wereldwijde standaard is ontwikkeld, gebruik makend van state-of-the-art methoden, voor de codering van videosignalen. MPEG kende twee fasen, namelijk MPEG-1 voor het ontwikkelen van een coderingsstandaard voor opslagmedia met een bit rate tot 1.5 Mbs, en MPEG-2 voor een meer algemeen toepasbare coderingsstandaard tot grofweg 30 Mbs.

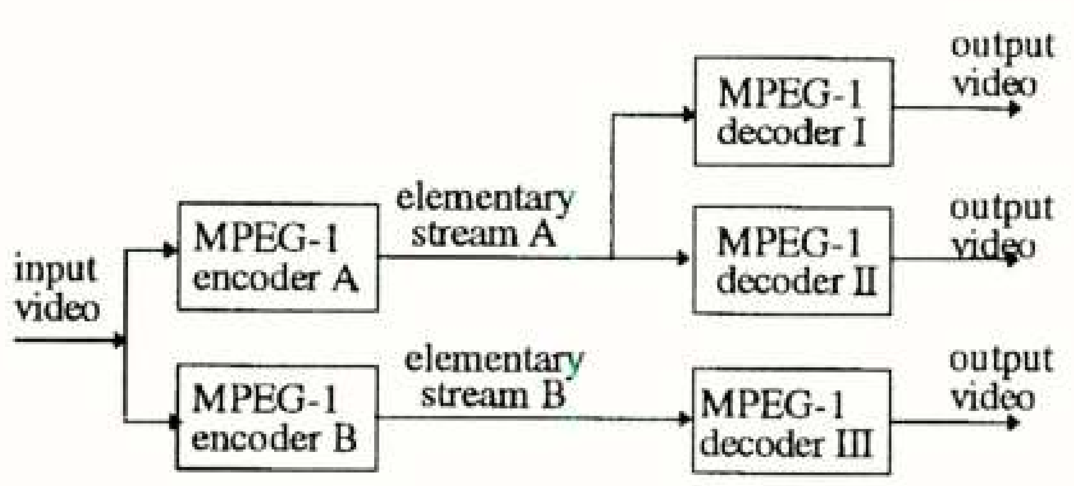
Beide MPEG standaarden [8,9] werken volgens het basisschema en de principes zoals in het voorgaande geïntroduceerd, zonder expliciet voor te schrijven hoe de codering van het videosignaal zou moeten worden uitgevoerd. MPEG legt namelijk alleen de *syntax* van de te transporteren bitstroom vast. Het maakt daarbij gebruik van een gelaagde syntax die direkt gekoppeld is aan de opbouw van een videosignaal. Figuur 11 toont deze gelaagde syntax. Hierin is zichtbaar dat een *videosignaal* opgebouwd gedacht wordt uit *GOPs*, waarbij elk GOP bestaat uit (I/P/B) *frames*. Elk frame bestaat uit een aantal *slices*, wat tevens de kleinste herkenbare eenheid in de MPEG syntax is. Zowel GOPs, frames en slices kunnen herkend worden in de bitstroom aan de hand van unieke *startcodes*. Een slice bestaat uit *macroblocks*, en elk macroblock bestaat uit 4 8x8 luminantie, en 2 8x8 chrominantie subbeelden die de basis vormen voor de bewegingscompensatie en DCT transformatie.

Het slechts vastleggen van de syntax van de bitstroom die het gecodeerde videosignaal representeerd, heeft als voordeel dat er een grote vrijheid is in het ontwerpen van een MPEG coder. Voor een gegeven videosignaal bestaan er dan ook talloze verschillende MPEG gecodeerde versies. De feitelijk uitgevoerde codering, gemaakte keuzen, opdeling in GOPS, I/P/B frames, slices, bewegingscompensatie etc. kunnen en zullen waarschijnlijk verschillen tussen verschillende MPEG coders, evenals natuurlijk de kwaliteit van het resultaat. Toch kan een videosignaal dat volgens MPEG gecodeerd is altijd door een willekeurige MPEG decoder gedecodeerd worden dankzij de nauwkeurig omschreven regels waaraan een MPEG bitstroom moet voldoen (zie Figuur 12)



Figuur 11: MPEG gelaagde syntax.

De belangrijkste verschillen tussen MPEG-1 en MPEG-2 zijn de volgende. MPEG-1 is grofweg geoptimaliseerd voor het coderen van *progressief* videomateriaal van *SIF formaat* bij bit rates rond de *1.2 Mbs*. MPEG-2 daarentegen is minder sterk geoptimaliseerd voor een bepaalde bit rate (5-15 Mbs), laat *geïnterliniëerde* videosignalen toe, en gaat uit van signalen in *4:2:0 YUV formaat*. Tevens laat MPEG-2 een nauwkeurigere representatie van de bewegingsvectoren toe dan MPEG-1.



Figuur 12: Twee verschillende MPEG coders leveren een verschillende MPEG elementary bitstream *op, die beide door willekeurige MPEG decoders gedecodeerd kunnen worden. De output van decoder I en II zijn identiek, terwijl decoder III een andere output oplevert.*

In zekere zin is MPEG-2 een *toolkit* benadering. De MPEG-2 syntax laat een groot aantal opties toe, maar voor een gegeven applicatie zal lang niet van alle mogelijkheden gebruik gemaakt worden. Enkele van deze opties zijn (i) het formaat van het te coderen videosignaal, (ii) de transmissie bit rates (iii) schaalbaarheid van de bitstroom en van het gedecodeerde videoformaat (iv) MPEG-1 compatibiliteit, (v) verhulling van transmissiefouten. Om enige orde aan te brengen in de veelheid aan opties zijn er MPEG *hardware profile* en *video levels* gedefiniëerd. Tabel III geeft een kort overzicht van de profiles en levels, waarbij de opmerking gemaakt kan worden dat momenteel de meeste MPEG-2 coder en decoders zich richten op het main profile en main level (main@main).

High	1920×1152	×	< 80 Mbs	×	< 100 Mbs
High-1440	1440×1152	×	< 60 Mbs	< 60 Mbs	< 80 Mbs
Main	720×576	< 15 Mbs	< 15 Mbs	< 15 Mbs	< 20 Mbs
Low	352×288	×	< 4 Mbs	< 4 Mbs	×
Level		Simple 4:2:0 no B not scalable	Main 4:2:0 not scalable	Next 4:2:0 scalable	High 4:2:2 scalable
		Profile			

Tabel III: Profiles en levels in MPEG-2.

5. Toekomstige ontwikkelingen

MPEG-1 en MPEG-2 zijn compressiestandaarden voor digitale video die momenteel hun weg aan het vinden zijn in een groot aantal toepassingen waarbij visuele informatie gecodeerd moet worden met goede kwaliteit bij bandbreedten van grofweg 300 kbs tot 30 Mbs. Het MPEG systeem zal zich ontwikkelen tot een bouwsteen in vele digitale televisie en multimedia applicaties.

Echter, MPEG kent ook beperkingen die momenteel aanleiding zijn voor een verschuiving van de aandacht naar zogenaamde object- en modelgebaseerde coderingstechnieken [6]. In de eerste plaats, wanneer een MPEG coder wordt toegepast bij lage tot extreem lage bitsnelheden, blijkt dat er inherente kwaliteitsbeperkingen zijn aan de opbouw van de MPEG syntax en aan de gelaagde structuur die een 8x8 subbeeld als kleinste eenheid kent. Ten tweede, de inhoud van een MPEG gecodeerde videosignaal is slechts toegankelijk nadat een geheel frame gedecodeerd is. Voor sommige toepassingen lijkt het van belang ook direct toegang te hebben tot delen van een frame zonder dat volledige decodering noodzakelijk is. Aan beide bezwaren kan tegemoet gekomen worden wanneer bij de codering niet langer uitgegaan wordt van (relatief arbitraire) 8x8 subbeelden, maar van betekenisvolle "objecten". Deze objecten vormen dan de basiseenheden waarop de bewegingsschatting, (DCT) transformatie, kwantisatie, etc. uitgevoerd zou moeten worden. Het vastleggen van wat betekenisvolle objecten zijn op zodanige wijze dat ook de codering daarvan efficiënt kan verlopen, is echter geen triviale vraagstelling. Dit vergt niet alleen de ontwikkeling van nieuwe generatie beeld- en video*compressietechnieken*, maar vereist tevens een grote inbreng vanuit het gebied van de beeld- en video*analyse*.

Referenties

[1] Jayant and P. Noll, *Coding of Waveforms*, Prentice Hall, Englewood Cliffs, NJ, 1984.

[2] M. Rabbani and P.W. Jones, *Digital Image Compression Technique*, SPIE Optical Engineering Press, Bellingham, 1991.

[3] R. Veldhuis and M. Breeuwer, *An Introduction to Source Coding*, Prentice Hall, New York, 1993.

[4] J.C.A. van der Lubbe, *Information Theory*, Cambridge University Press, Cambridge, MA, 1995

[5] D. LeGall, "The MPEG video ompression algorithm", *Signal Processing: Image Communication*, vol. 4. pp. 129-140, April 1992.

[6] H. Mussman, P. Pirsch and H. Grallert, "Advances in picture coding", *Proc. IEEE*, vol. 73, pp. 523-547, April 1985.

[7] E.D. Frimout and J. Biemond, "Video compression: Techniques and standardization activities", *Proc. of the IEEE Symp. on Multimedia*, pp. 13-40, Delft, October 1992.

[8] ISO/IEC International standard ISO/IEC 11172: Coding of moving pictures and associated audio for digital storage media up to about 1,5 Mbit/s.

[9] ISO/IEC International standard ISO/IEC 13818-2: Generic coding of moving pictures and associated audio - part 2: video.

Voor een meer diepgaande studie van beeld- en videocodering wordt verwezen naar de jaarlijkse 5- daagse PATO cursus " Digitale Video: Coderingstechnieken en toepassingen van visuele communicatie".

Voordracht gehouden tijdens de 431e werkvergadering

NEDERLANDS ELEKTRONICA- EN RADIOGENOOTSCHAP
WERKGEMEENSCHAP INFORMATIE- EN COMMUNICATIE-THEORIE
IEEE BENELUX CHAPTER ON INFORMATION THEORY
IEEE BENELUX CHAPTER ON COMMUNICATIONS AND VEHICULAR TECHNOLOGY
431e werkvergadering

UITNODIGING voor de gezamenlijke werkvergadering van het IEEE, WIC en NERG op donderdag 12 januari 1995 in Lecture Hall WB van Philips Research, Roland Holstlaan4, Waalre

THEMA: Digital Television Broadcasting (DVB)

Inspanningen op het gebied van het digitaliseren van videosignalen heeft geleid tot een doorbraak op het gebied van het toepassen van coderingstechnieken voor televisietoepassingen. Als alternatief voor het huidige PAL systeem zullen deze technieken op grote schaal toegepast gaan worden in de vorm van digitale TV via kabel, satelliet en aardse zenders. Deze werkvergadering gaat in op de laatste stand van zaken op dit gebied.

PROGRAMMA:

09.50 uur	Welcome **DR.IR. T.A.C. CLAASEN** (Director Philips Research)
10.00 uur	Standardization by the European Project on Digital Video Broadcasting (DVB)
	IR. F.W.P. VREESWIJK (Philips Research)
10.30 uur	The Principles of (MPEG) Video Coding
	DR. R. LAGENDIJK (TUD)
	Coffee
11.30 uur	MPEG/DAVIC; Interactive Services
	IR. J. A. KOSTER (PTT Research)
12.00 uur	MPEG Multiplexing and ATM Transmission
	IR. R. TER HORST (PTT Research)
	Demonstrations, lunch
14.00 uur	Satellite and Cable Broadcasting
	IR. L.A. VERVOORT (Philips CE)
14.30 uur	Terrestrial Broadcasting
	IR. P.G.M. DE BOT (Philips Research)
	Coffee
15.30 uur	Conditional Access for Pay TV
	IR. D. VAN SCHOONEVELD (Philips Research)
16.00 uur	Digital Television on CATV Networks
	IR. F.J.W. VAN LET (CASEMA)
16.30 uur	Closure **PROF.DR. E.C. VAN DER MEULEN** (Chairman WIC)

The working language will be English

Aanmelding voor deze dag dient te geschieden vóór 6 januari aanstaande door middel van de aangehechte kaart, gefrankeerd met een postzegel van 70 cent.

De lunch wordt aangeboden door Philips Research. Routebeschrijving zie achterzijde.

IEEE, WIC en NERG leden hebben gratis toegang. Het maximaal aantal deelnemers is 125.
Indien u door overinschrijving niet kunt deelnemen, ontvangt u vooraf bericht.

Organisation: Dolf Schinkel (070-3325006; dagvoorzitter)
Wim van der Bijl (070-3325744; programma manager NERG)
Cees Jansen (040-723497; contactpersoon WIC)
Paul de Bot (040-742702; contactpersoon IEEE)

A BRIEF INTRODUCTION TO DAVIC
THE DIGITAL AUDIO-VISUAL COUNCIL

R. Koenen, A. Koster
KPN Research

DAVIC'S vision

The Davic forum was conceived in January '94, at the initiative of Dr. Leonardo Chiariglione (CSELT), the convenor of MPEG. All participants in DAVIC share a vision of a digital audio-visual world, where producers of digital audio-visual content can reach the widest possible audience, users have seamless access, carriers can offer effective transport, and manufacturers can provide hardware- and software to support unrestricted production, flow and use of information. The first meeting was attended by a group of people from 17 countries and 40 companies and organisations representing the majority of players in the digital audio-visual field. DAVIC held an establishment meeting on 2nd and 3rd June , in San Jose, CA, USA where about 150 poeple representing 96 companies participated.

Promoting audio-visual services

The purpose of DAVIC is the promotion of the success of the emerging digital audio-visual applications and services, first for broadcast and interactive use. As these services are intended for home use, cost plays an important role. DAVIC believes that these services can only be made affordable if sufficient standardisation ensures reasonable prices through the possibility of mass production, and fair competion.

Many technical questions still need an answer. By combining efforts, DAVIC can answer these questions in a shorter period of time, thus reducing the time to market.

Lastly, DAVIC wants to prevent customer confusion. The existence of many different, incompatible kinds of home equipment ('set top boxes') on the market, will create puzzled and hesitating consumers. And, consequently, a slow start for service providers.

DAVIC seeks to ensure the smooth introduction of the new interactive services, by promoting the timely availability of internationally agreed specifications of open interfaces and protocols. The objective is to maximise interoperability across countries and services, through the open international collaboration of all players in the field. The concept of "Digital Audio-Visual Applications and Services" is taken to include all those applications and services in which there is a prevalent digital video component. The DAVIC Statues prescribe that DAVIC results shall be made available to all interested parties, and will be contributed them to the appropriate international standards bodies.

Using what is available

DAVIC wants to make as much use of available specifications as possible, be they of interfaces, protocols and architectures. The task is 'just' adopting the right-ones. Unfortunately, not all necessary elements exist yet, so some specifications will have to be augmented, and some others will even have to be developed from scratch. The order reflects the desirability of the way specifications should be produced by DAVIC.

In working toward the achievement of these stated purposes and objectives, the Council and its members declare to be individually and collectively committed to open competition in the development of digital audio-visual products, technology and services, and the members are not restricted in any way from designing, developing, marketing and/or procuring digital audio-visual hardware, software, systems, technology, or services. Implementation or use of specific digital audio-visual standards, recommendations and DAVIC specifications will be voluntary, and no member shall agree or be obliged to implement them, just because they are members of DAVIC.

Relationship with international Standard Bodies

DAVIC aknowledges the importance of the various international standards bodies and is keen to co-operate with them. In particular, DAVIC expects to utilise elements of existing standards, augmenting or developing them as necessary. Furthermore, the results of activitieswithin DAVIC, such as specifications, will be made available to all interested parties and will be contributed to the appropriate international standards bodies.

Structure of the Technical Committees.

At its fourth meeting the environment for the technical work was created. A total of 5 Technical Committees were set up:

* *Set-Top Unit* will specify the 'reference set top unit' and its interfaces;

* *Server* will specify all the Video server interfaces, but nothing from the inside of the 'black box';

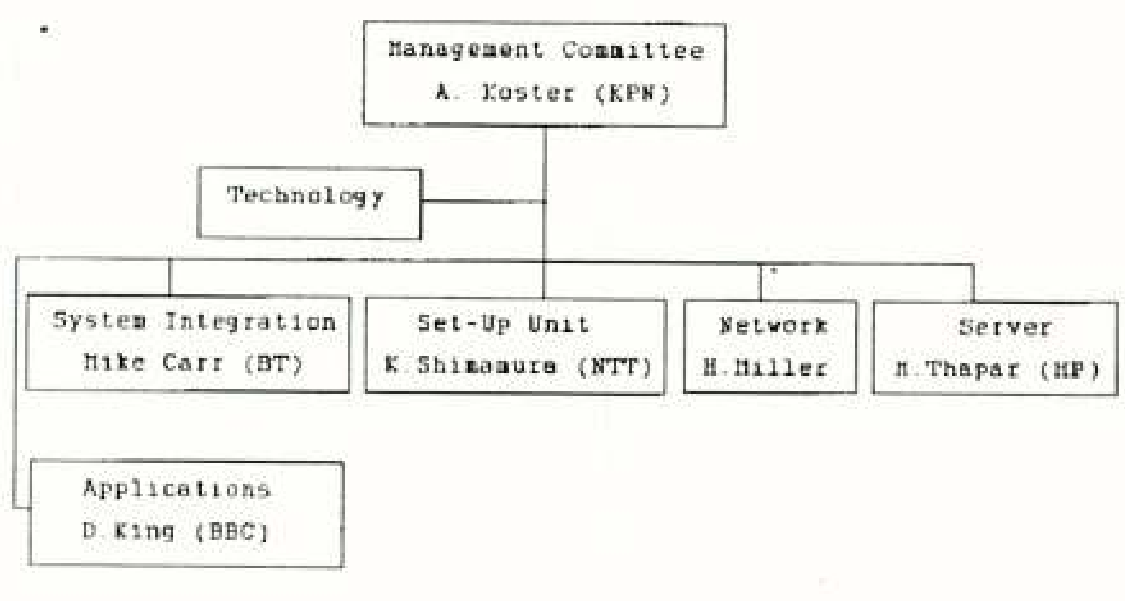
* *Network* will specify a 'reference model' for each delivery system and all the interfaces, and will specify delivery methods across networks;

* *Systems Integration Applications* will specify the 'reference model', deal with the overall system issue,

* *Systems Integration* handles definition of DAVIC Core services and service issues like copyright management, broker functions, application portability etc;

* *Technology* will keep inform DAVIC on technological and standard issues;

The structure and relations of these committees is shown in the diagram below. The current workplan of DAVIC is the production of specification of those



interfaces and protocols that are needed by the so-called DAVIC core services, e.g. Video on Demand.

The Technical Committees, under the supervision of the Management

Committee, are currently producing the text of a Call for Proposals that will be widely distributed. Based on the responses obtained DAVIC will produce and issue the planned spacifications by 1st December 1995.

For further inquiries please contact:
Dr. Leonardo Chiariglione
Multimedia and Video Services
CSELT
Via G. Reiss Romoli, 274
10148 Torino
ITALY
Tel.: +39 11 228 6120
Fax: +39 11 228 6299
Email: leonardo.chiariglione@cselt.stet.it

Voordracht gehouden tijdens de 431e werkvergadering

DIGITAL TERRESTRIAL TELEVISION BROADCASTING

Paul G.M. de Bot
Philips Research Laboratories

Abstract

In the coming years, the current analog television distribution will be replaced by digital distribution. Standards for digital transmission via satellite and cable have been developed for this purpose, and a standard for digital terrestrial is on its way. In this paper, the technical details of digital terrestrial television broadcasting will be described.



Introduction

Recently, practical systems for video source coding have been developed in the framework of the ISO/MPEG project. This effort has lead to a growing interest for introduction in Europe of digital broadcasting services in the near future. With this respect, we should distinguish between satellite direct-to-home distribution, cable network distribution and terrestrial distribution. Since these distribution media each have different channel characteristics

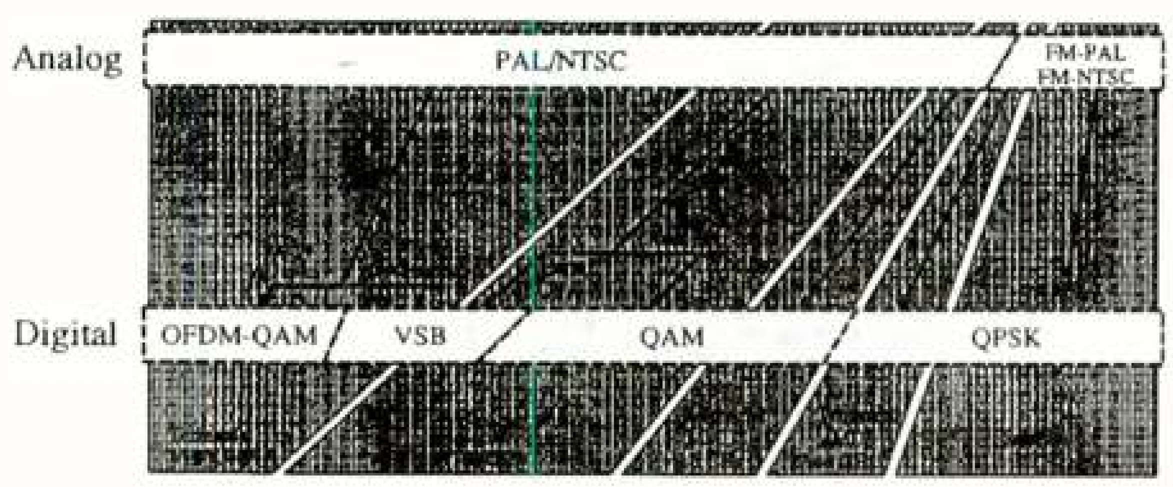


Figure 1: Migration from Analog to Digital for Different Television

and require different receiver equipment, different transmission mechanisms have to be designed, each optimized for a specific medium. All of these mechanisms enable the transport 24-40 Mbit/s in a single channel. Since MPEG-2 source coding can provide good standard definition video quality at bit rates of 4-8 Mbit/s, such a transport stream is sufficiently large to contain a number (4-8) of normal standard definition TV programs. In Figure 1, migration from analog transmission to digital transmission is depicted for the different television transmission media. A generic picture of a digital television chain is given in Figure 2.

Early 1994, a draft European standard has been fixed, describing a transmission mechanism for TV broadcasting via satellite [1].

Satellite transmission is characterized by low available transmitter power, relatively high channel bandwidth (33-40 MHz), highly nonlinear transmitter amplification and a transmission channel which approaches the Additive White Gaussian Noise (AWGN) channel.For these reasons, QPSK modulation is chosen with powerful concatenated error correction coding.

For cable TV networks, another transmission standard is drafted this year [2]. The cable channel is characterized by a high signal-to-noise ratio, a strong bandwidth limitation (8 MHz), and short reflections due to impedance mismatches in the network. These constraints have lead to the choice of 64-QAM modulation and interleaving in combination with a single Reed-Solomon code. For compatibility reasons, the interleaving and Reed-Solomon coding are chosen the same as for the satellite system.

The terrestrial channel is for sure the worst and most difficult of the three

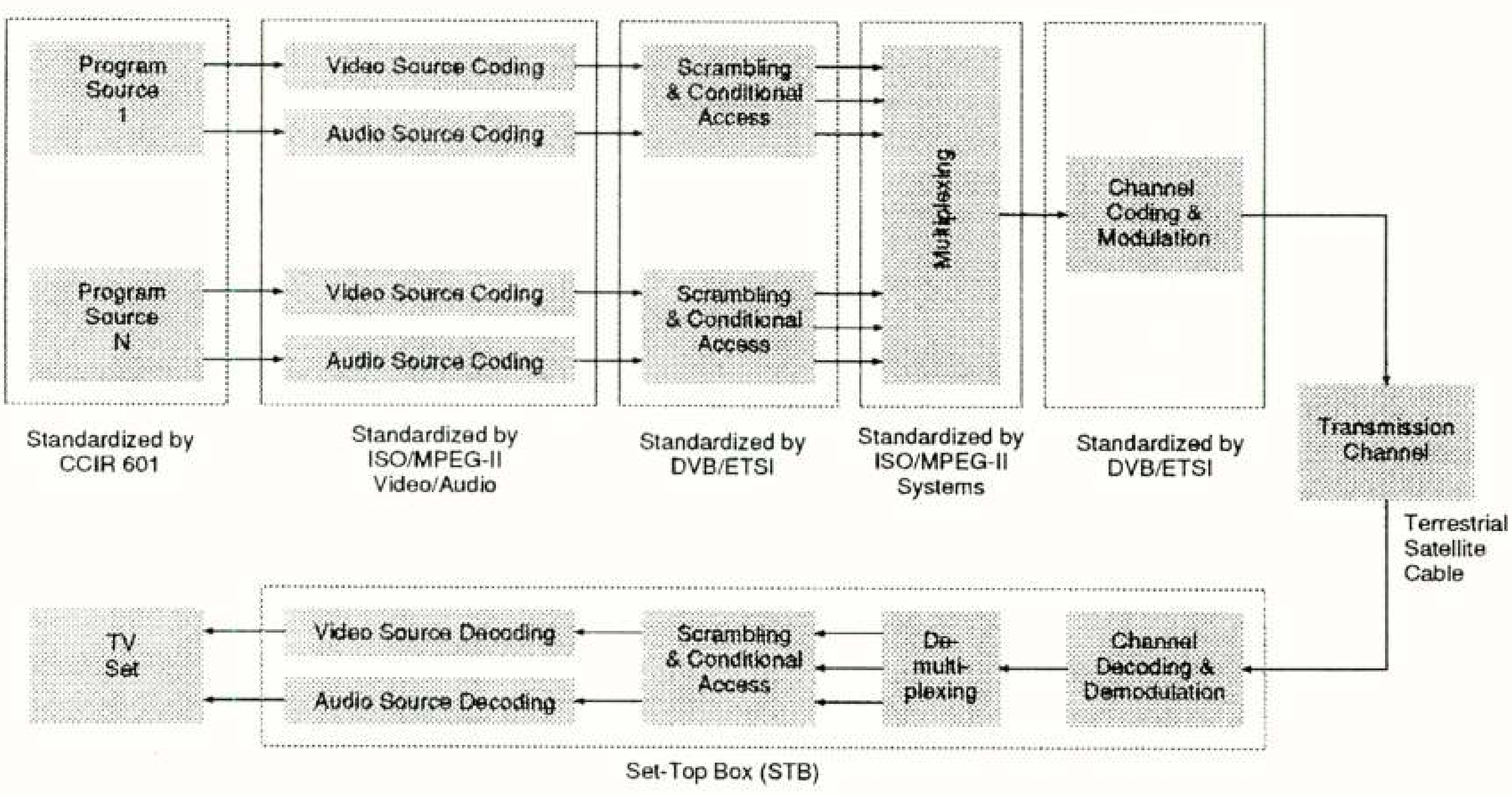


Figure 2: The Digital Television Chain

channels discussed. For this reason, no final standard has yet been fixed in Europe. Discussions in all three European DTTB projects[1] focus on the use of Orthogonal Frequency Division Multiplexing (OFDM), in contrast to the single carrier systems chosen for satellite and cable. Also in Japan, OFDM-based systems are considered for Digital Terrestrial Television Broadcasting (DTTB), although in North America a Single Carrier VSB solution is chosen. In the rest of this paper, OFDM-based DTTB system will be described in more detail.

2 Channel Characteristics for Terrestrial Broadcasting

The DTTB system should allow large coverage for fixed receivers (with a directional roof-top antenna), and also provide the largest possible coverage for portable receivers (indoor reception with a non-directional built-in antenna). However, these two reception conditions are related to different transmission channels. Fixed reception coverage will be mostly interference limited, where the interferer, in the DTTB introduction period, probably is a PAL/SECAM signal. The transmission channel for portable reception however, will mainly be characterized by multipath propagation, resulting in a frequency selective, noise limited channel. Single Frequency Networks (SFNs), as will be described in Section 9, cause an effect, similar to multipath propagation, also to fixed receivers.

Conventionally, network planning is based on fixed reception. In [6], it is shown that transmitters should increase their power with something in the order 30 dB to offer the same service area for portable receivers. Therefore, introduction of DTTB services will be focussed on fixed reception.

On noisy multipath channels, the received signal r can be modeled as $r=\alpha_M(f)s+n$, where $\alpha_M(f)$ is the complex attenuation factor of the channel, s is the transmitted signal with $E[ss^*]=E_s$, and n is a complex additive white Gaussian noise (AWGN) component with $E[nn^*]=N_0$. Usually, $\alpha_M(f)$ is complex Gaussian distributed, and frequency dependant. Now we can define the frequency dependant signal-to-noise ratio $\gamma_N(f)=|\alpha_M(f)s/n|^2 = |\alpha_M(f)|^2 E_s/N_0$.

A typical profile for $\gamma_N(f)$ is shown in Figure 3 (left).

If a channel suffers from CCI that has a complex Gaussian amplitude distribution, the received signal can similarly be written as $r=s+\beta(f)n$, where $\beta(f)n$ represents the combined CCI and AWGN. If the CCI is caused by PAL/SECAM signals, β will be heavily frequency dependant and show power concentrations near the luminance, chrominance and sound carriers of the PAL/SECAM signal.

As well as for the case of multipath propagation, we can define the frequency dependant signal-to-noise ratio (where in this case the 'noise' is in fact interference) as $\gamma_I(f) = |s/\beta(f)\, n|^2 = |\alpha_I(f)|^2 E_s/N_0$, where $\alpha_I(f)=1/\beta(f)$. A typical $\gamma_I(f)$ profile, for a channel suffering from CCI of PAL is shown in Figure 3 (right).

Hence, we can deal with CCI from PAL/SECAM in the same way as with multipath propagation. Channels suffering from CCI from PAL/SECAM as well as channels with multipath propagation will be referred to as frequency selective channels. In the following, we will restrict ourselves to the term signal-to-noise ratio (SNR), although we mean signal-to-noise & interference ratio.

3 Orthogonal Frequency Division Multiplexing

In section 2 multipath transmission has been discussed, which is a predominant factors limiting the service coverage for terrestrial broadcasting. Assume that we have a channel allowing the transmission of symbols with duration T_s If the channel generates echoes, the effect is limited if the echo delay τ is small compared to T_s. To improve the resistance to echoes further, the symbol of duration T_s can be extended with a so-called guard interval of length T_g, containing a cyclic continuation of the same symbol. This yields symbols of total duration T_s+T_g, and reduces the transmission efficiency of the channel.

If an echo occurs with delay $\tau < T_g$, the received symbol shows overlaps with the previous symbol and the next symbol. However, in the center of the period T_s+T_g, a window of width T_s can be found, which is not corrupted by intersymbol interference (ISI). If a receiver is able to properly position observation windows of length T_s appropriately over the received signal, the transmitted symbols can be recovered without suffering from ISI.

However, in SFNs (See Section 9), the echo delay can be as large as 200 μs. This means that the guard interval should have a duration of $T_g=200\ \mu s$. To ensure a sufficiently large transmission efficiency, the (Nyquist) symbol period should be chosen not smaller than T_s=800 μs, yielding an efficiency loss of 20 % due to the guard intervals. If we transmit the symbols with a rectangular pulse shape in the time domain, the Fourier transform of the signal $s(t)$ will be $S(f)=T_s\ sinc\ ((f\text{-}f_c)Ts)$, where f_c is the frequency of the carrier. If T_s=1 ms, the effective bandwidth of the signal is F_s=1 kHz. Since channels of 8 MHz are available for DTTB, we could combine many such narrowband signals in the wideband transmission channel. If we use signals $s_k(t)$ with carrier frequencies $f_{c,k}$ of each exactly F_s=1 kHz apart, the signals are orthogonal. This means that at the receiver side, the different signals $s_k(t)$ can be recovered without any mutual cross-talk. This technique is known as Orthogonal Frequency Division Multiplexing (OFDM). OFDM isproposed for DTTB transmission in Europe and Japan. OFDM is also being used in the Digital Audio Broadcasting (DAB) system [7] [8].

To combine the many narrow-band signals into a wide-band OFDM signal, an Inverse Discrete Fourier Transform (IDFT) can be used at the transmitter side, combined with a DFT at the receiver side. By using a complex IDFT of N= 8192 points, we can multiplex N signals $s_k(t)$ with k=0,....,N-1onto an 8 MHz channel. In Europe, this so-called 8K OFDM scheme is proposed for DTTB.

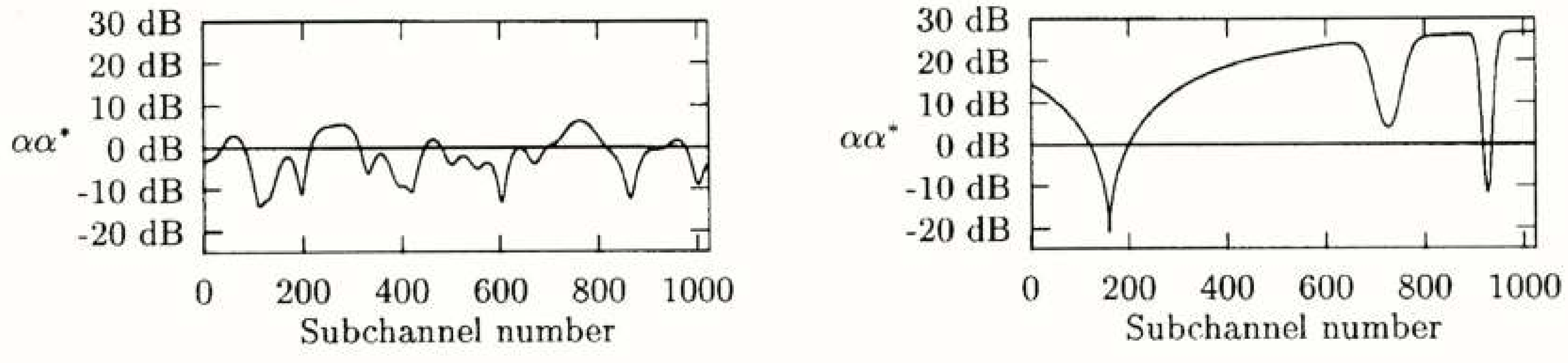


Figure 3: $\gamma n(f)$ profile of multipath reception in an urban area (left) and $\gamma I(f)$ profile of PAL Co-Channel Interference (right). The wideband value of γ equals 0 dB in both cases.

[1] *The RACE dTTb project [3], the German HDTVT project [4] and the Nordic Divine project [5]*

Since guard bands in the frequency domain are needed for filtering, a number of carriers at the edges of the 8 MHz channel are modulated with a zero-signal. Effectively, the wideband signal gets in this way a bandwidth of some 7.5 MHz. If the carriers are 1 kHz apart, this means that we effectively modulate $N_{eff} \approx 7500$ carriers.

Hence, in each OFDM time slot (with a duration of T_s+T_g), we can transmit N_{eff} complex symbols. These symbols can be PSK or QAM symbols, as will be described in Section 4. The total gross symbol rate over a channel equals $N_{eff}/(T_s+T_g)$. An example of an OFDM scheme is given in table 1 .

FFT size	N	8192
sampling frequency	B	9.1429 MHz
symbol period	$T_s = N/B$	896 μs
carrier spacing	$F_s = 1/T_s$	1.17 kHz
guard interval	$T_g = T_s/4$	224 μs
effective number of subchannels	N_{eff}	6785
effective signal bandwidth	$B_{\text{eff}} = N_{\text{eff}} F_s$	7.57 MHz
gross symbol rate	$R_g = N_{\text{eff}}/(T_s + T_g)$	6.06 Mbaud
net symbol rate	$R_n = R_g(93/96)(7/8)$	5.14 Mbaud

Table1: Example of an OFDM scheme for nation-wide SFNs

By using OFDM with guard intervals, the problem of ISI in the time domain is solved. However, the frequency selective nature of the channel (due to both multipath and CCI) causes each of the OFDM carriers k to be subject to to a different signal-to-noise ratio $\gamma k = \gamma(f_{c,k})$. Error correction coding is needed to recover the information transmitted on the carriers which are subject to low γ values. Error correction coding will be described further in Section 5.

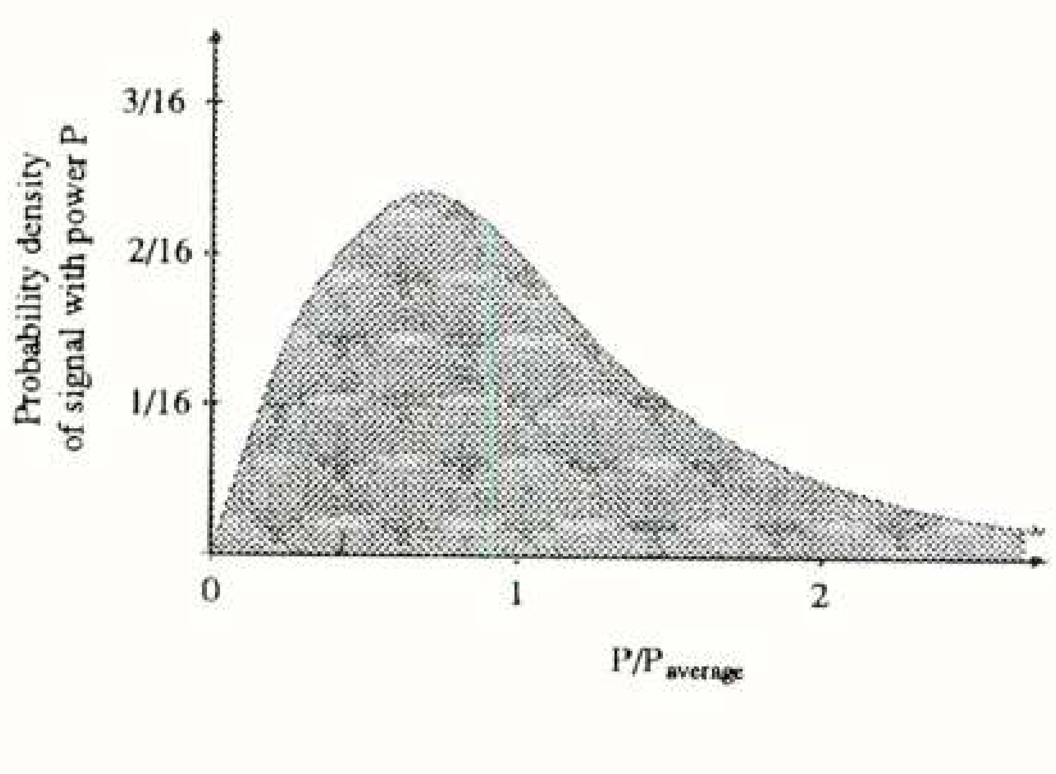


Figure 4: Power Distribution of OFDM (large N)

In addition to error correction, the frequency selectivity can be reduced by using antenna diversity with narrow-band combining. This technique can improve the performances on severe frequency selective channels with up to 10 dB [9] [10].

The output of the IDFT at the transmitter side, is a signal of which the amplitude has a very high dynamic range, and a very high peak-to-average power ratio is very large (See Figure 4). Therefore, clipping will inevitably occur at the high power transmitter amplifiers, due to its non-linear behavior. The percentage of samples which is corrupted by this clipping, will depend on the output back-off of the power amplifier. The larger the percentage of samples which is clipped, the more the performance at the receiver side is degraded. Since broadcasters want to operate with cost-effective power amplifiers, a trade-off has to be determined between the nominal amplifier power (and the used output back-off), and the acceptable degradation due to clipping. In Figure 5, the DTTB transmission chain is shown.

4 Modulation

As explained in Section 3, we can modulate one complex symbol in each timeslot on each useful carrier. These symbols are elements of a symbol set. If we put all the symbols of a symbol set in a complex plane, we obtain the so-called signal constellation. Typically, a signal constellation contains $M=2^m$ signal points, which means that each symbol carries m bits of information. Hence, M should be large to obtain a large transmission rate. On the other hand if M is large, the required signal-to-noise ratio to obtain a desired error rate, is also large. For choosing a signal constellation, we have to make the trade-off between transmission rate and required signal-to-noise ratio.

For DTTB, usually 3 modulation schemes are considered; 4-PSK (or 4-QAM), 16-QAM and 64-QAM, with M=4, M=16 and M=64, respectively. To obtain a sufficiently low error rate for these modulation schemes we need signal-to-noise ratios in the order of E_s/N_0=6 dB, E_s/N_0=12 dB and E_s/N_0=18 dB, respectively. The signal constellations with added noise of the critical SNR values, are shown in Figures 6-8.

The m bits will be mapped on the signal points in a signal constellation using Gray mapping. In this case, an error event will cause a minimum number of bit errors.

Some DTTB proposals foresee the option of hierarchical transmission. In this case, the modulation and error protection are organized such, that at the receiver side, different bit streams can be extracted from the received signal, each with a different a priori reliability. For example, we can use non-uniform QAM signal constellations, transmitting m_{HP} high priority (HP) bits and m_{LP} low priority (LP) bits per symbol. On bad channels, a receiver will only be able to recover the HP bits reliably, while on a good transmission channel, also the LP bits can be detected with a low error probability. Examples of hierarchical DTTB systems are described in [11], [12].

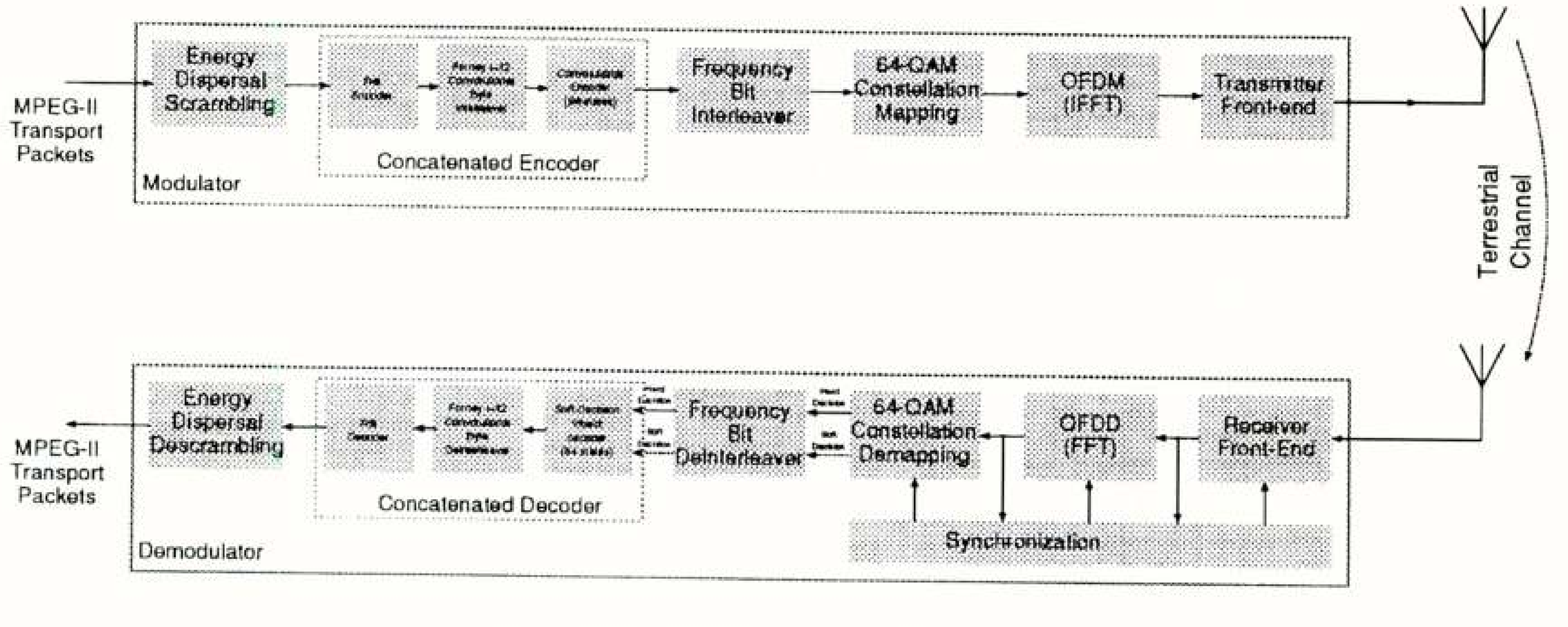


Figure 5: The DTTB Transmission Chain

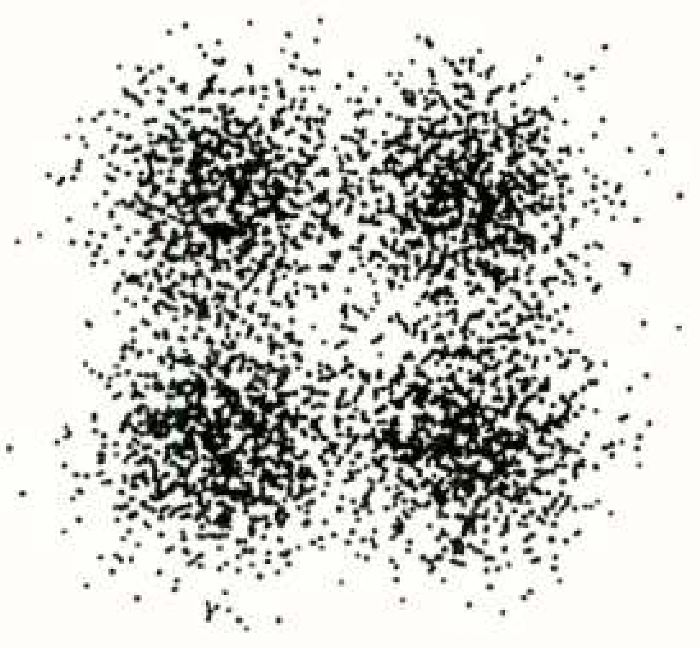


Figure 6: 4- QUAM Signal Constellation with AWGN (E_s/N_o=6 Db)

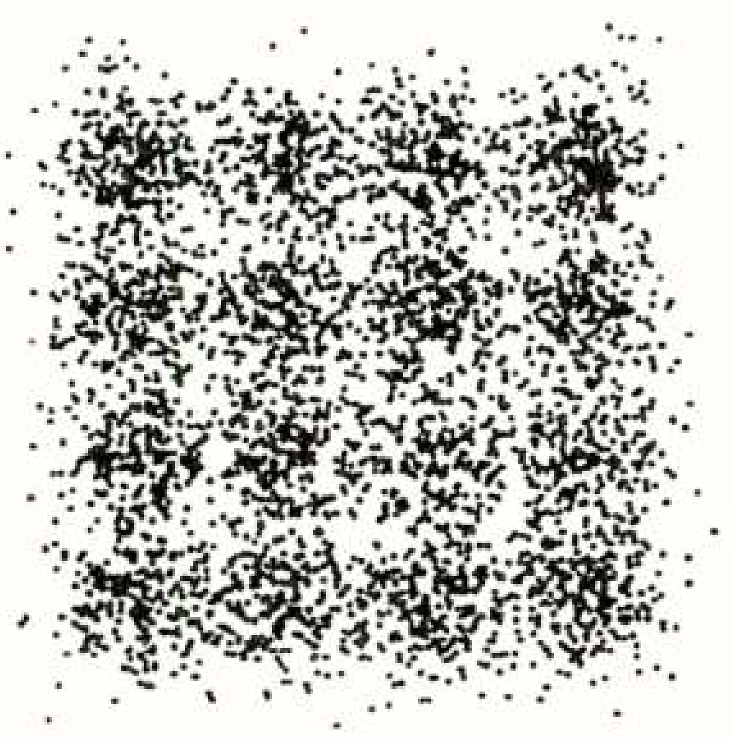


Figure 7: 16 -QUAM Signal Constellation with AWGN (E_s/N_o = 12 dB)

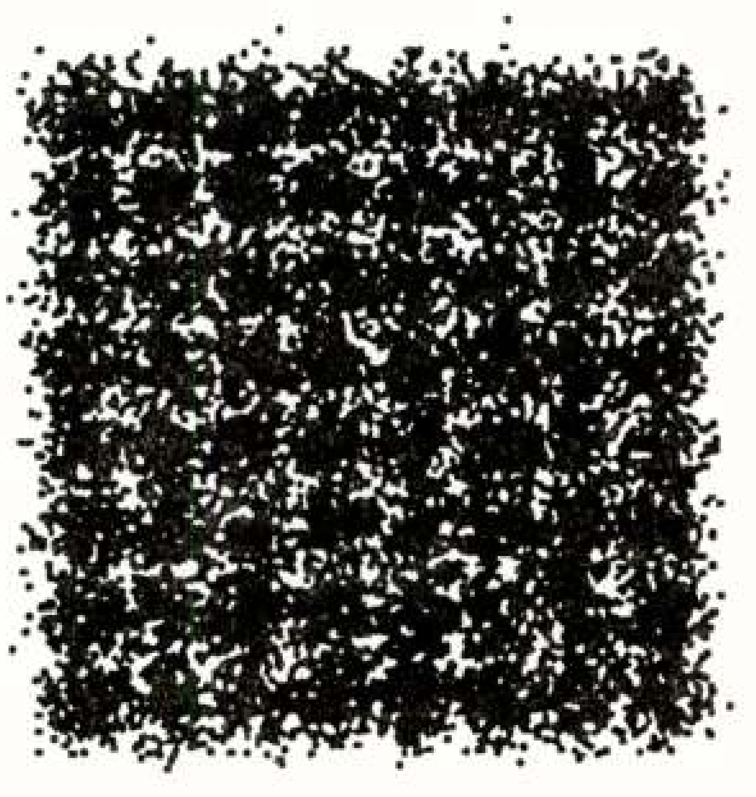


Figure 8: 64 -QUAM Signal Constellation with AWGN (E_s/N_o = 18 dB)

5 Error Correcting Coding

Error correcting coding is used to guarantee at the input of the demultiplexer virtual error-free performances (i.e. a bit error rate (BER) of 10^{-10}). To achieve this, the same concatenated error correcting codes is used, as in the digital satellite television standard. This concatenated code consists of a v=6 64-state (punctured) convolutional inner code with code rates R=1/2, 2/3, 3/4, 5/6 and 7/8, a Reed-Solomon [204,188,17] code over GF(2^8) and appropriate interleaving.

The inner convolutional code has the advantage of being decodable with the Viterbi algorithm, which is an implementable Maximum Likelihood (ML) decoding algorithm. This makes the code a powerful tool to reduce a BER from 10^{-1}-10^{-2} on the channel to 10^{-3}-10^{-4}. Since the Viterbi decoder is only able to correct random bit errors, and has no burst error correction capabilities, inner bit interleaving in the frequency domain is applied, to scatter the frequency selective behavior of the channel and provide random bit errors at the input of the Viterbi decoder. In the Viterbi decoder is not able to correct certain errors, it will, depending on the puncture rate of the code and of the channel) typically produce burst errors of 5v-15v(equivalent to 30-90) bits. Since the outer RS-decoder is able to correct random byte errors (but not bursts of byte errors), the bits at the Viterbi decoder output are organized in bytes, on which outer byte interleaving is applied using a Forney interleaver of a depth of 12 bytes. Thanks to its large Hamming distance d=17, which makes correcting possible of up to 8 random byte errors per codeword, the Reed-Solomon decoder is able to reduce the bit error rate further e.g. from 10^{-3}-10^{-4} down to10^{-10}-10^{-11}.

The Viterbi algorithm can be implemented for either hard-decision or soft-decision decoding. Soft-decision decoding, relative to hard-decision decoding, of a v=6, R=1/2 convolutional code with BPSK modulation can yield a gain of up to 4 dB in the region of BER=10^{-4} on a Rayleigh fading channel [13]. Because a soft-decision Viterbi decoder uses $r=\alpha s+n$ at its input, the channel state α is implicitly being used by the decoder. Explicit knowledge of α, by the availability of Channel State Information (CSI), can yield some additional gain. In case of frequency-selectivity, the optimal method for combining this CSI with the received signal, is by applying maximal ratio combining [14]. Using this method, we will supply $\alpha^* r$ to the input of the Viterbi decoder. Hence, strong signals are made stronger, while weak signals are made even weaker. The use of CSI in this way, can yield an additional gain of 2 dB in the region of BER=10^{-4} on a Rayleigh fading channel [13]. It is furthermore interesting to see that if we use hard-decision Viterbi decoding with 1-bit CSI(comparable with an erasure flag), we closely approach the performances of soft-decision Viterbi decoding without CSI [13].

If the channel is not primarily subject to multipath, but to interference, it is essential that this interference is being estimated in order to apply soft-decision decoding.

Methods for estimating the CSI are given in Section 7.

6 Synchronization

Good reception of the transmitted signals is only possible if good frequency and time synchronization is achieved. To enhance synchronization, the FFT blocks, or OFDM symbols, are organized in frames. In the European DVB project, a frame containing a number of 96 OFDM symbols is proposed, including a silent period (null symbol). This null symbol is used for coarse time synchronization. Immediately after the null symbol, a reference symbol is being transmitted that has good autocorrelation properties in the frequency domain. This reference symbol is used for coarse frequency synchronization. Since the transmitted content of the reference symbols is known a priori at the receiver side, the receiver is able to estimate the frequency domain transfer function $H(f)$. Hence, the receiver can calculate the time domain transfer function $h(t)$, and is consequently able to determine the optimal observation window timing (fine time synchronization). The reference symbol on its turn is followed by a Transmission Parameter Signalling (TPS) symbol, which contains transmission mode information, such as the used signal constellation and convolutional code rate. The TPS symbol is modulated and protected such that it can be received even under very bad channel conditions.

Since oscillator stability of the receiver can be a limiting factor, the frequency synchronization needs to be extremely accurate. A small frequency error causes a fixed rate of phase rotation in each QAM signal, or cell, as well as cross-talk between the subcarriers. To support more accurate frequency synchronization, the remaining 93 symbols in a frame contain a certain amount of pilots, which do not contain data and from which the receiver can estimate the actual frequency offset. In order to minimize the cross-talk, this frequency error signal is being fed back and compensated for prior to the FFT. Since

OFDM is very sensitive towards frequency jitter and phase noise, the local oscillator in the receiver front-end needs to have a very high level of frequency accuracy.

In the frame structure described above, 93 of the 96 OFDM symbols in a frame can be used for data transmission, while in these 93 symbols, about 12.5 % of the carriers are used to transmit pilot symbols. Hence, the net symbol rate equals about (93/96) (7/8) times the gross symbol rate. For the example of Table 1, this yields a symbol rate of 5.14 Mbaud.

7 Channel Estimation

Furthermore, channel amplitude and phase estimation are required for enabling coherent detection of the QAM signals, and for providing reliability information to the soft-decision Viterbi decoders. This channel estimation is performed by evaluating the reference symbol and pilot cells. To obtain information about the frequency domain characteristics of the co-channel interference (CCI) the null symbol is evaluated. More accurate estimation of the S/(I+N) ratio in each subcarrier can be obtained by examining the statistics of the received signal [15].

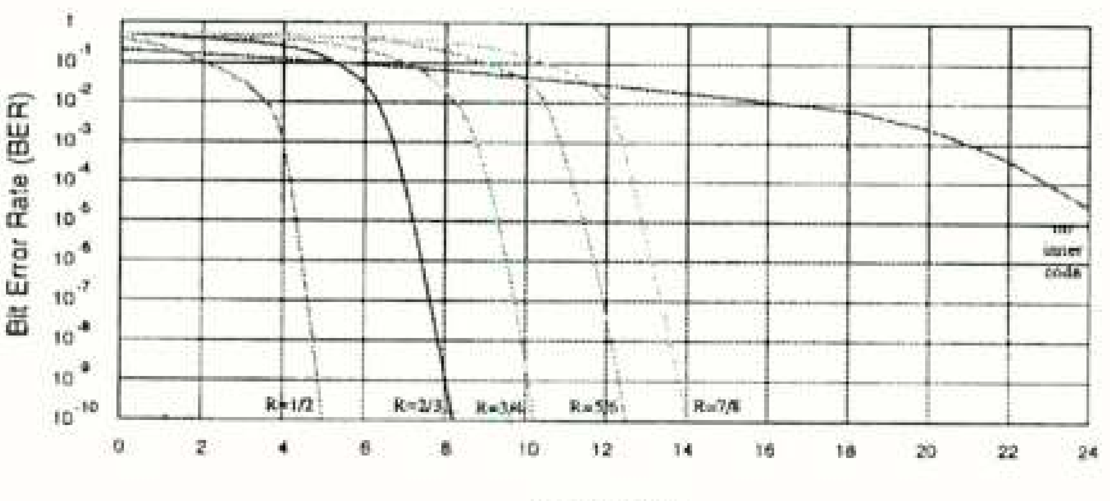


Figure 9: Perfomance of 4-QAM with concatenated coding on a Rayleigh channel, for different inner code rates R.

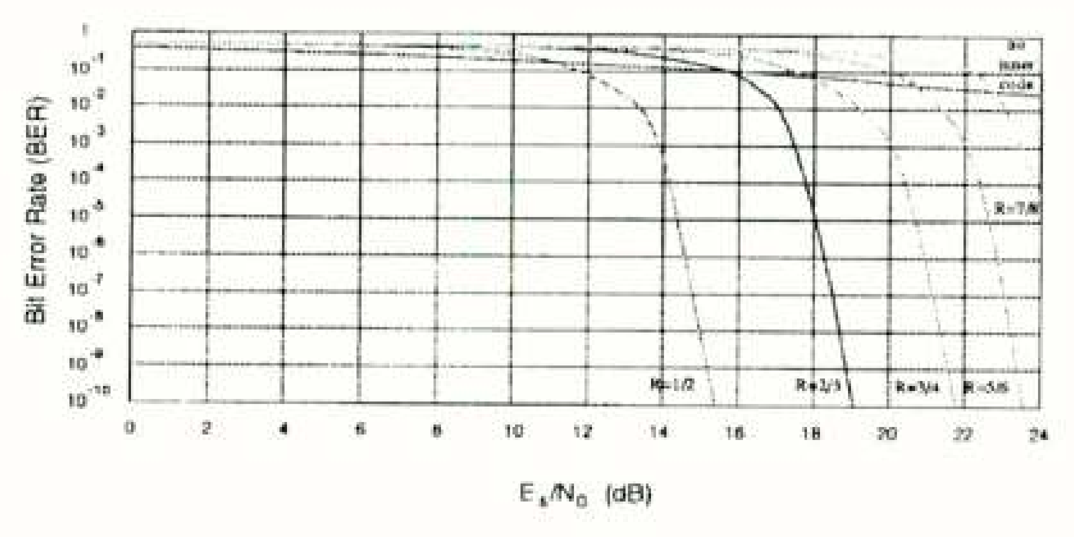


Figure 10: Performance of 64-QAM with concatenated coding on a Raleigh channel, for different inner code rates R.

8 System Performances

In Figures 9 and 10, the BER curves are given for a few modulation/code combinations on a Rayleigh channel. It can be seen, that the required SNR ranges from less than 5 to more than 25 dB, depending on the transmission mode.

Furthermore, we should notice that in case of CCI from PAL (See Figure 3) the receiver will give sufficiently low error rates for signal-to-interference ratios (SIR) as low as 0 dB (depending on modulation/coding).

This remarkable result can be explained by the fact that even with an SIR of 0 dB, most of the OFDM carriers are subject to a narrow-band SIR of more than 20 dB, since the interfering power is concentrated on only a limited number of OFDM carriers. The error correcting codes are easily capable of

	QPSK	16QAM	64QAM
$R = 1/2$	4.70	9.40	14.10
$R = 2/3$	6.27	12.53	18.80
$R = 3/4$	7.05	14.10	21.15
$R = 5/6$	7.83	15.67	23.50
$R = 7/8$	8.23	16.45	24.68

Table 2: Net Bit Rate (in Mbit/s) of a typical DTTB System (See Table 1), for different modulation schemes and inner convolutional code rates R.

correcting the errors made in this limited number of interfered OFDM carriers. Its ruggedness towards CCI from analog services, is one of the major advantages of OFDM.

The net bit rate can be calculated from the net symbol rate, and depends on the modulation and coding. It is given in Table 2.

9 Single Frequency Networks

A Single Frequency Network (SFN) is a broadcast transmitter network consisting of transmitters with overlapping coverage area's that transmit the same program in the same frequency channel at the same time instant. Consequently, the same signal can arrive at a receiver antenna from different SFN transmitters, each with its own delay, related to the distance between receiver and transmitter. The receiver can deal with this effect in the same way as it deals with multipath propagation (the signals arriving from distant transmitters are considered as echoes from the signal arriving from the nearby transmitter).

Since conventional analog transmission schemes (as PAL television) cannot cope with extreme multipath, SFNs were traditionally not possible. However, since OFDM systems with guard intervals are inherently capable of handling multipath, SFNs become practical.

Since SFNs improve the efficiency of spectrum usage considerably, the SFN-feature is an important advantage of OFDM systems over analog and single carrier digital systems[2]

We can distinguish between local SFNs, consisting of a single main transmitter and a few gap-fillers to cover shielded areas, and nation-wide SFNs, which consist of a large number of main transmitters.

Distribution of the signal from the central studio to the main transmitters can take place in various ways, using micro-wave or optical fiber links, satellite feeding, or by mutual in-channel feeding of the transmitters throughout the network. Synchronization of the transmitters in an SFN is still an issue of intensive study.

In nation-wide SFNs, the delay spread can be as large of T_m=200μs, causing the need for an guard interval of T_g=200μs and an 8K DFT, as described in Section 3. As an alternative for using a large Tg and large DFT, SFN echoes can be cancelled using mixed time/frequency-domain equalization [16].

10 Service Introduction

Many difficulties have to be overcome, before DTTB can be introduced in Europe.

Receiver ICs still have to be designed, SFN network structures have to established, including the transmitter synchronization, commercially viable introduction scenarios have to be developed, and channel space has to be allocated.

According to an optimistic scenario, the first introduction of DTTB is planned around 1998. The United Kingdom has already made available channel space for service introduction. It is the intention in the UK to establish a nation-

[2] *Digital single carrier systems would require an extremely long adaptive equalizer at the receiver side, which is very complex, compared to an OFDM receiver.}.*

wide SFN in UHF channel 35, while throughout the country, in each region at least 2 taboo channels are assigned for regional and local-SFN operation. In the London area, it is even expected that 7 channels of 8 MHz each can be made available. If on average 4 TV programs can be accommodated in each channel, the London area will have access to 28 DTTB programs.

Other countries which could have some perspectives for service introduction are Denmark, Sweden and The Netherlands. In countries like Germany and Italy, all parts of the broadcast spectrum are completely filled with analog TV services. Due to this lack of available channels, introduction of DTTB in these countries seems to be very difficult.

11 Conclusions

The opportunities for the introduction of DTTB are far from clear. However, in Europe a technical solution has been developed, albeit at the cost of high receiver complexity. During a period of at least 10-15 years after service introduction, simulcast with analogue PAL/SECAM will be required, which enlarges the frequency allocation problem. After switching off the analogservices, DTTB with SFNs will improve the spectral efficiency of television broadcasting significantly. This means that on the long term, broadcast spectrum could be made available for non-broadcast applications.

References

[1] ETSI. Channel coding and modulation for ll12 GHz satellite receivers, April 1994. prETS 300xxx/6.

[2] ETSI. Framing structure, channel coding and modulation for CATV cable and smatv distribution, May 1994. prETS 300xxx/7.

[3] P.G.M. de Bot, B. Le Floch, V. Mignone, and H.D. Schütte. An overview of the modulation and channel coding schemes developed for digital terrestrial television broadcasting within the dTTb project. In Proc. Int. Broadcasting Convention, pages 569-576, Amsterdam, The Netherlands, September 1994.

[4] R. Schäfer and J. De Lameillieure. MPEG coding of TV and HDTV for an integrated satellite, cable and terrestrial broadcast system. In Proc. Int. Broadcasting Convention, pages 321- 326, Amsterdam, The Netherlands, September 1994.

[5] S. Nystrom, V. Ringset, G. Roth, and E. Stare. A flexible COFDM modem for the HD-Divine digital terrestrial broadcasting system. In Proc. Int. Broadcasting Convention, pages 585-588, Amsterdam, The Netherlands, September 1994.

[6] P.G.M. de Bot, A.J.M. Wijlaars, H. Vermeer, and H. Kollenbrander. Propagation of broadcast signals into buildings at 762 mhz. In Proc. *IEEE* Symp. on Communications and Vehicular Technology in the Benelux, pages 28-37, November 1994.

[7] M. Alard and R. Lassalle. Principles of modulation and channel coding for digital broadcasting for mobile receivers. EBU Review, (224):168190, August 1987.

[8] D. Pommier and Y. Wu. Interleaving or spectrum-spreading in digital radio intended for vehicles. *EBU* Review, (217):128-144, June 1986.

[9] P.G.M. de Bot. Antenna diversity for OFDM systems. In Proc. Symp. on Information Theory in the Benelux, pages 244-251, Veldhoven, The Netherlands, May 1993.

[10] J.G.W.M. Janssen, P.G.M. de Bot, and A.J.M. Wijlaars. Antenna diversity for digital video broadcasting. In Proc. *IEEE* Symp. on Communications and Vehicular Technology in the Benelux, Delft, The Netherlands, October 1993. (dTTb/WP3.2/Philips/24).

[11] K.M. Uz, K. Ramchandran, and M. Vetterli, Multiresolution source and channel coding for digital broadcast of HDTV. In *Proc. HDTV Workshop,* Turin, Italy, September 1991.

[12] P.G.M. de Bot, C.P.M.J. Baggen, A. Chouly, and A. Brajal. An example of a multi-resolution digital terrestrial tv modem. In Proc. Int. Conference on Communications, pages 1785-1790, Geneva, Switzerland, May 1993.

[13] J. Hagenauer. Viterbi doding of convolutional codes for fading- and burst-channels. In *Proc. 1980 Zürich* Seminar on Digital Communications, pages G2.1-G2.7, Zürich, Switzerland, 1980.

[14] D. G. Brennan. Linear diversity combining techniques. *Proc. IRE,* 47(6):1075-1102, June 1959.

[15] P.G.M. de Bot and C.P.M.J. Baggen. Reed -Solomon codes for OFDM broadcasting over frequency selective channels using error & erasure decoding. In International Tirrenia *Workshop* on digital Communications, pages 203-214, Tirrenia, Italy, September 1993.

[16] A.G.C. Koppelaar and C.P.M.J. Baggen. Twodimensional equalization for OFDM systems. In Proc. Symp. on Information Theory in the Benelux, Veldhoven, The Netherlands, May 1993.

Voordracht gehouden tijdens de 431e werkvergadering

DIGITAL TELEVISION ON CATV NETWORKS

Ir. F.J.W. van Let
Manager Technical Department of the N.V. CASEMA The Netherlands

Summary

Digital Video Broadcasting (DVB) as a system is very interesting and attractive for the operators of transmission-infrastructures. It offers new opportunities because it provides additional methods of use of the available frequency space on their networks and therefore new possibilities to add new services and extra incomes. These facts also apply to the operators of Cable Television Systems (CATV). In a heavily cabled country like The Netherlands the successful introduction of DVB even depends to a great extent on the behaviour of the CATV networks and that of their operators.

This article describes the situation of the Dutch Cable market with respect to DVB from a technical- as well as a service point of view. It explains how these very new techniques will be implemented in the CATV networks that CASEMA has built since the mid-seventies and that have been upgraded periodically to remain prepared for the future. It also indicates which important things with respect to DVB on Cable still have to be done.



Introduction

Undoubtedly most of you have a connection to a CATV network. Therefore most of you belong to the broad public that is watching to the ever growing number of television programs delivered by Cable on that very simple wall-outlet. You are zapping more or less frequently through the mere 15 to 25 programs for which you have to pay a moderate monthly fee. For a few special programs however you need a special decoder and you have to pay additionally for them. The number of so called pay-television-programs however will grow significantly in the years to come. Also Pay-Per-View programs will become increasingly available. That will be achieved by digitizing the television programs on the CATV networks.

Before I explain how we at CASEMA have planned to do so, I will give a short historic overview with respect to the developments in CATV networks during the past years and how we are already upgrading our networks for our customers, the Cable viewers, for the digital times to come.

Historic overview

Cable television has become a permanent feature of life in The Netherlands. In The Netherlands 5.8 million houses are past by the Cable and 5.5 million of them are actually connected to the networks. Currently 5.2 million , that is 90 %, are actual subscribers. CASEMA was set up in 1970 by the NOZEMA, the Dutch PTT and several broadcasting companies. Its objective was to create a nationwide cable television network, but for political reasons it did not go as planned. Despite that at present CASEMA has about 1.2 million subscribers in 113 towns in The Netherlands.

When CASEMA started, in 1970, Cable television was still in its infancy. At that time the network operators were relatively amateurish. All they wanted was a shared facility to replace aerials on flats. They improvised and got whatever hardware they could lay their hands on. They were quite happy if their systems could carry six television channels with a reasonable quality. In those days it was quite sufficient. There is a world of difference between this approach and today's professional cable TV systems. Before long, the operators introduced a structured approach to the installation of networks. The cable TV networks had to be expanded as the increasing number of high flats blocked transmissions. Furthermore the larger networks required a much better structure than before. The first CATV concept used by CASEMA in the seventies was the so called C18 system. This was still based on the hardware available from the industry at that time. However, there was one major difference between practice in The Netherlands and elsewhere in the the world: the regulations forced us to use the same frequencies on the Cable as were used for terrestrial transmissions. The problem associated with this and the reason why the rest of the world did not follow this approach, was that the availability of hardware for such high frequencies on Cable was very limited. The C18 system developed at that time was based on reconversion, a distinction was made between the primary trunk-network and the district networks. The primary networks contained amplifiers with a frequency range up to 300 MHz. In a district centre the signals of the primary network were converted to frequencies up to 860 MHz. At any given time cable TV in The Netherlands was more advanced than in most other countries. The networks at that time , the late '70 's, could carry up to 20 television channels, albeit with some difficulty and by using special frequency grids. This was partly due to limitations imposed by the consumer's TV-sets. For example: it was not possible to use adjacent channels.

These 20 channels were sufficient for a long period, until satellite made its entrance in the mid-eighty's. Operators in The Netherlands were suddenly faced with a major increase in the number of TV-channels in a short time. It did not take them long to realise that their cable networks, limited to 20 channels, would not last very long. Something had to be done! At that time the hardware available on the market did not meet CASEMA's requirements. However the hybrids -the major components in cableTV amplifiers- had undergone major technical improvements. On the basis of its experience with cable television networks built up over the years CASEMA decided to have its own equipment manufactured as it was not available on the market. The systemconcept that should make use of this new generation of amplifiers was called C30, that described the network architecture and included a list of technical requirements of the network elements. The C30 concept was based on the assumption that the active elements of the network would be used for at least 10 years and the passive part would be used for 20 to 30 years. All amplifier designs were based on a broadband, twoway technique. Although there were no twoway- or interactive applications at that time there were possible applications expected to be introduced eventually. Besides that it was of course also an attractive feature for PR purposes. A positive effect for the industry was, that they sold even more "CASEMA amplifiers" abroad than in The Netherlands, whereas CASEMA itself bought more than 30.000 examples!

The capacity of the current C30 networks is limited to about 30 television channels. One of the limiting factors affecting the total capacity of those cable networks is that the channel spacing was based on the assumption that five HDMAC 12 MHz broad channels and a number of digital audio channels would be implemented. However, at present it is most unlikely that techniques such as HDMAC and DSR (Digital Satellite Radio) will ever be successful. After all, the C30 system died an early death. However, it was designed to allow for future expansion. The technical and economical lifespan of the network are quite different. The untimely demise of the system was due to the fact that the demand for capacity grew much more rapidly than expected. A number of new services which require very large bandwidths have been developed and others are still in the stage of development. Examples include Pay-Per-View, datacommunications and telephony. Besides that, at that time - in 1992- the price of glassfibre equipment was already falling rapidly. Previously, the cost of glassfibre systems had always been prohibitive.

As a result a new system, called C50, was developed to upgrade to glassfibre in the CATV networks. The architecture of the CASEMA glassfibre networks is fundamentally different to that of the coaxial networks currently in use. Technically was decided to implement a "staged" star network. That means, that the network starts conventionally at a central point and branches to the distribution centres. From such a distribution centre the laser can feed 3 to 4 district centres as the distances to be covered are quite limited. For the time being, glassfibre will only be used for the main trunk network. If the capacity needs to be expanded (due to the introduction of new -interactive- services) while using the same district network, the quality of the trunk network in front of that, will have to be improved considerably. For this reason the glassfibre trunk networks are based on an old principle: split band. Two fibres will be installed: one for distribution of the band of 45 to 450 MHz (modulated on an optical carrier of course) and one for 470 to 862 MHz. A third fibre will be installed for the return path. CASEMA will install additional fibres for the business data- and communications market, so there will be integration of networks rather in cables than in fibres.

Picture quality on demand

All these activities relate to upgrade for growth, CASEMA's main objective. To date an expansion of the capacity of the CATV networks could only be realised by making more bandwidth available for the subscribers. The required bandwidth per television program was always a constant factor.

However an operator of an infrastructure will try to achieve an optimised return on investments on his infrastructure. That can be realised -amongst other possibilities- by an optimised choice of transmission techniques. The factor "money per hertz" is of increasing importance and in a technical sense for that reason also the factor "bits per hertz". In fact it makes no difference whether it concerns a satellite operator or a cable operator, despite the fact that the resulting technical choices will differ. When the information-density rises and -as expected- the costs of transmission will be reduced, a rising number of new services will be possible. As a result again the better the portfolio of cableservices, the better the operational and financial results of the cable TV network will be.

Returning to some technical matters: To distribute an uncompressed televisionsignal (HF-PAL) including the associated stereophonic sound over the cablenetwork, a bandwidth of about 6 MHz is required. A normal base-band video-signal has a bandwidth of 5 MHz. Transmitting such a signal in a digital manner over the CATV network -without any compression- with a modulationtechnique such as 64 QAM, will result in a HF bandwidth of about 19 MHz! By satellite, using QPSK, such an uncompressed signal would even require about 63 MHz.... (figure 1)

Clearly compression is an absolute must for digital signals on networks. However, using compression has even enabled multiplication of the capacity of each (CATV-)channel with a factor 4 without loss of the reputable current

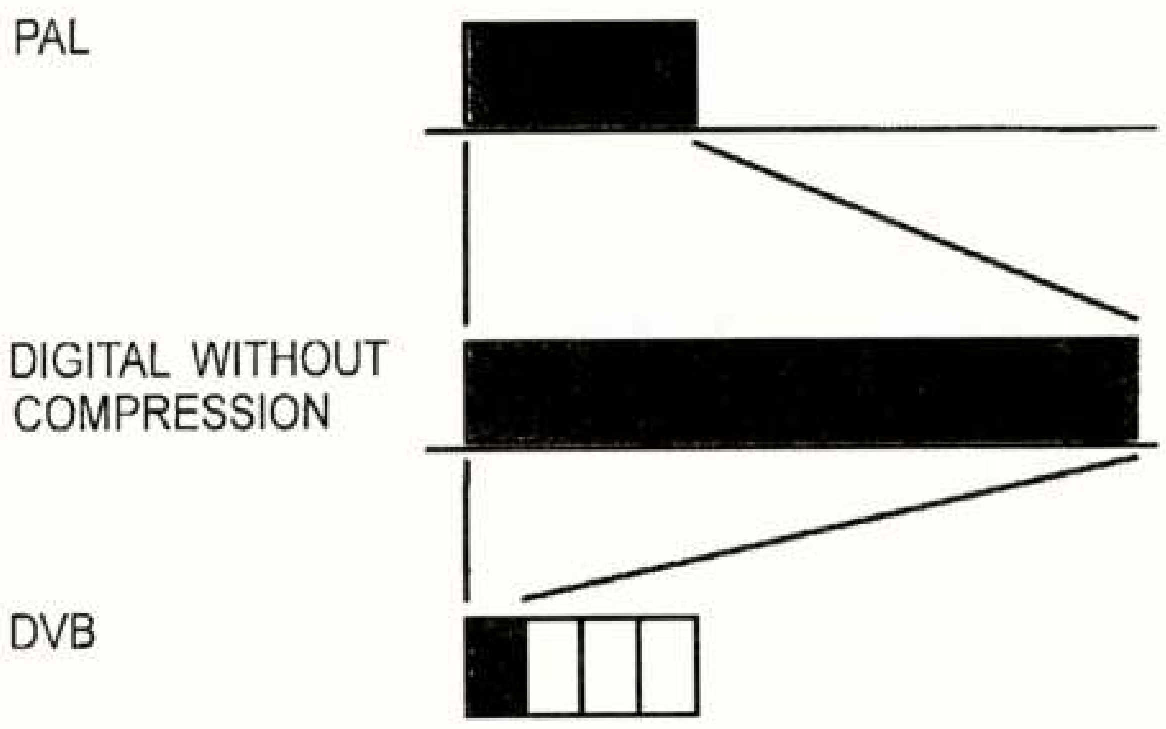


figure 1 Efficiënt compression + modulation: profits

quality of a PAL-signal. That is the first point of interest in "digital" for the operators of infrastructures. The second one is the the possibility to allocate more or less bits to each televisionprogram to be transmitted. That of course results in different properties of the channels (e.g. more or less sharpness of the picture) but also in different costs. Simply said: the transmission of television programs with fast moving pictures in high-definition quality will cost more (for programprovider and/or subscriber) than the transmission of a program with mainly still pictures, e.g. a cable-magazine.

But these facts and possibilities are the same for cable-operators and satellite-operators, the last one being the main competative operator in the near future for the cable operators. To date that was not really the case: one could even speak of some energetic effects with the satellite path and the cable path as a cascaded transmission chain. Accounting for the high penetration results for Cable in The Netherlands one might conclude that the price-performance relationship for reception via Cable (including matters as the ease of connection) is still better than for direct reception via satellite, though at the moment the number of channels available on a CATV network is already significantly smaller than on all the satellites together. However there are two important factors that are able to influence the price-performance relation and by that the choice of the customer:

a) only a few multi-channel-orbit positions come into being, which give individuals the possibility to quite easily receive a growing number of radio- and television programs with a simple fixed dish;

b) the introduction of DVB in Europe will lead to the situation that reception of the complete package of satelliteprograms can be realised with one standardized receiver only.

Therefore the main advantages of the present day Cable are rather slight. The competition between the two for the customer equivalent infrastructures, will be -if the cable operators do not change their policy- in the pricing area only. That will be a tough task. So it will be of eminent importance to look for services that will be unique for cable operations making use of the new possibilities of DVB techniques. Besides that it is very important that DVB-satellite-services can be transferred to signals on CATV networks without any constraint and without exceptional costs in the headends, distribution systems and receivers in the subscriber's homes. It must be said, that in principle an unconstraint translation of DVB signals from the satellite path to the cable path has been made possible. But principle and practice can differ considerably. As an example: within the DVB standard there are possibilities to use a fixed or a variable bitrate, even in a dynamical way. Choices will of course be made by the satellite program providers and - operators. Probably the cable operators will have to make a program-mix from different sources (e.g. satellite transponders) into one cable channel. That means, that the sum of the bitrates of the mixture can be variable and dynamic in time. One can imagine, that the sum of bitrates for periods of time could be higher than the maximum available

bitrate of the cable channel. There are several methods to prevent such a situation, but they will generally result in higher costs of head-end equipment and/or less bits/herz in periods of time and therefore loss of money.

Modulation

For an efficient distribution of digital videosignals via CATV systems only two methods of modulation are currently available: QAM (Quadrature Amplitude Modulation) and VSB (Vestigial Side-band Modulation). Both systems are particular forms of amplitude modulation.

Let us firstly look at QAM: In a transmission channel on one frequency two AM modulated signals can be distributed. The only thing is that the phases of the two carriers must have a difference of 90 degrees; they are -to say- in quadrature. Modulating each carrier in 8 levels -as is the case with 64 QAM- gives a total of 64 independant situations for the combined carrier that is modulated simultaneously in amplitude and phase (figure 2). Each situation represents a transport of 6 information-bits. That is by far not the maximum information density that could be possible on Cable as predicted by theory:

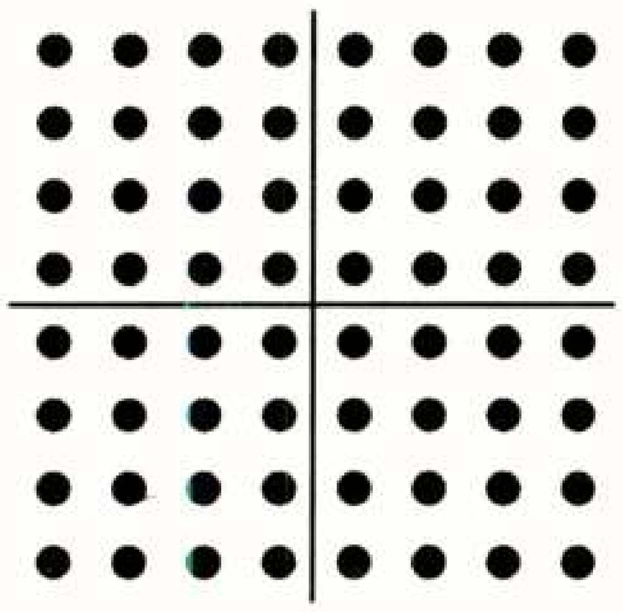


figure 2 64 QAM

Shannon says even 14 bits/s/hertz are possible.

It is possible of course to use less or more levels per carrier. That results in e.g. 16 QAM or 256 QAM. But the levels of the higher QAM forms are closer to each other, so the necessary carrier to noise distance must be better. Besides that, there are many other distortions present on a CATV systems and all of them have their influence on the signal. With special means one can try to reduce the effects; e.g. to prevent fatal influences of reflections in the CATV systems on the DVB signals in practice an equalizer must be built in every QAM receiver.

In the USA the choice has been made for VSB. VSB is also in use in Europe and other parts of the world for the well known analog PAL signals. The carrier is amplitude modulated and of course two side bands result containing the same information. One of them can be discarded (figure 3). In the digital area there is modulation only with a number of discrete amplitude levels. In fact the same capacity results as with QAM; e.g. 8-VSB has the same capacity

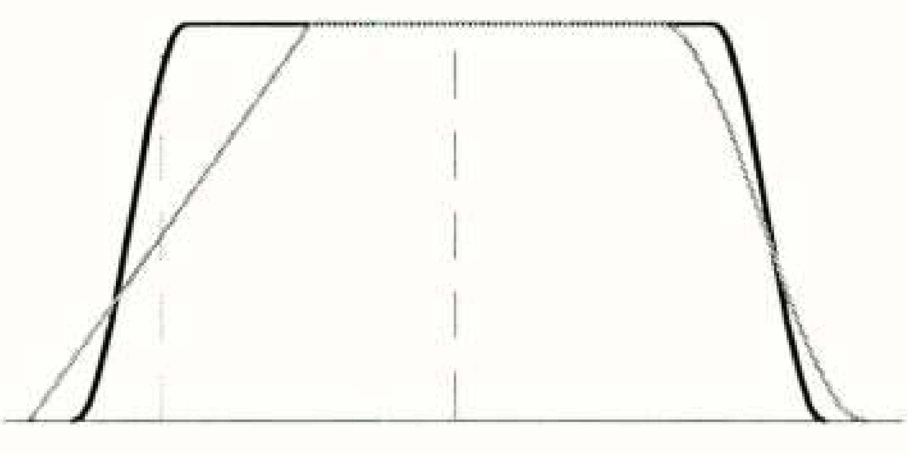


figure 3 Frequency spectra of VSB and 64 QUAM

as 64 QAM. Comparing QAM and VSB shows, that QAM is somewhat less sensitive to distortions such as phasenoise, jitter and groupdelay. Besides that in a VSB signal the carrier is always present; that carrier costs power and causes more intermodulation in the CATV network. But after all the choice for AM in Europe seems to be more political than technical.

Technical experiments in Holland

In the beginning of 1994 CASEMA was one of the first cable operators in Europe that performed tests with QAM signals in a normal large CATV system in operation. The provider of the at that time quite unique equipment was PHILIPS. The information obtained from these experiments has been offered to the DVB project via ECCA (European Cable Communications Association) and by CASEMA-people that participated in workinggroups themselves.

Based on theoretical studies we have taken a close look at HF-signal-to-noise, reflections, phasenoise and the protection curve for 64 QAM. For the moment we will take a look to the carrier to noise distance (C/N).

In figure 4 you can see the relation between the bit error rate (BER) of a 64

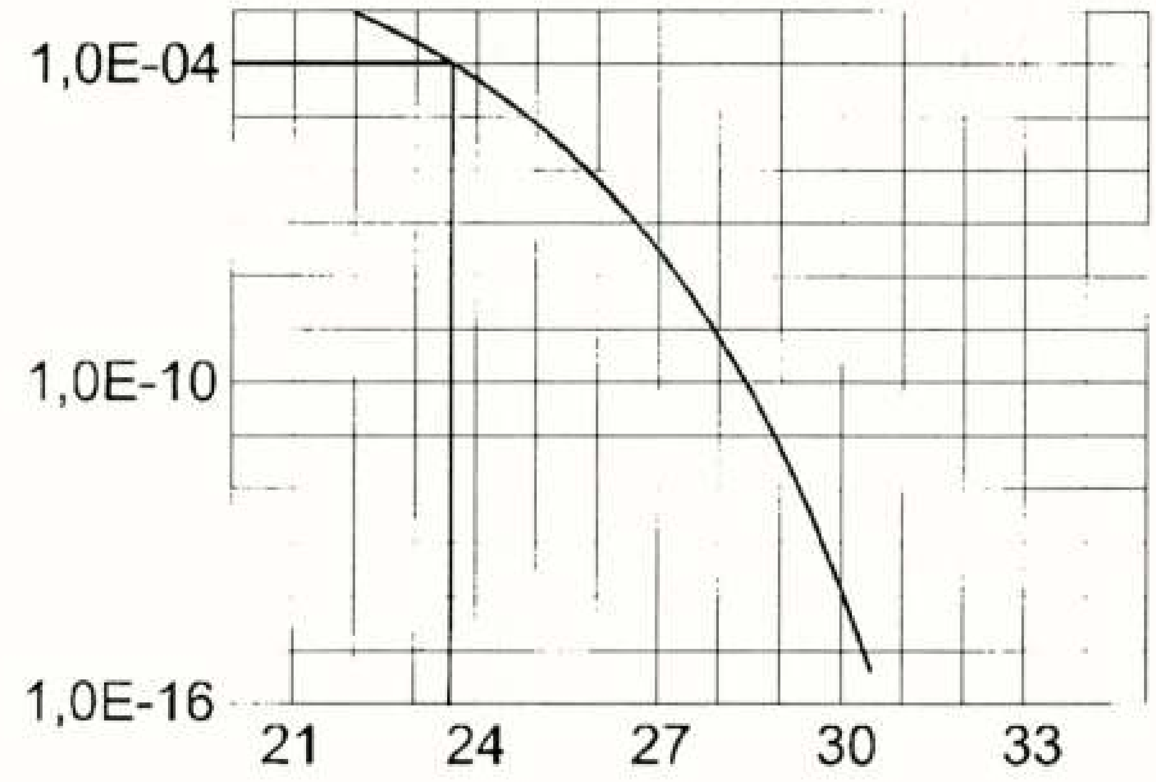


figure 4 BER versus C/N

QAM signal and the carrier to noise distance available, that can be reached on a theoretical base in a bandwidth of almost 8 MHz and a roll-off that is equal for modulator and demodulator. Because we have to reach a BER of 1 x 10E-4 for a quasi failure free DVB signal (after Reed Solomon decoding),

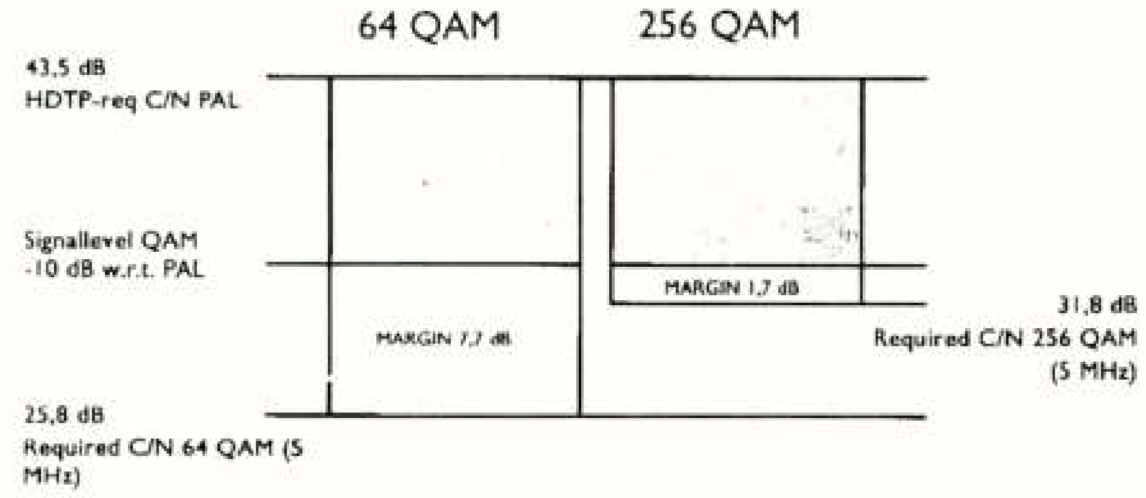


figure 5 QUAM in Dutch CATV- networks

figure 5 shows that a C/N of at least 23,6 dB is required.

To obtain a link with the Dutch Regulations for CATV systems we must recalculate this C/N to a bandwidth of 5 MHz. That offers 2,2 dB more, so 25,8 dB C/N results. The requirement for PAL signals on the wall-outlet at the subscriber's home is 43,5 dB and that offers a CATV systemmargin of 17,7 dB. That seems quite a lot, but we still have to consider intermodulation.

In order to reduce the systemload by these 64 QAM DVB signals we chose e.g. the level of these signals on -10 dB with respect to the PAL signallevels. In that case a systemmargin of almost 8 dB still exists (figure5). The same picture shows the schedule for 256 QAM. As you can see the margin to the noise floor is too slight to ensure a good transmission of such QAM signals: the signallevels for 256 QAM should be somewhat higher, -7 dB for example. During the practical measurements we did not reach the theoretically precalculated values. That was mainly due to the experimental character of the equipment at that time. Yet the results were very useful. In general terms the main conclusions were, that those CATV networks that fulfil the requirements will not encounter many difficulties. However the in-house systems of the subscribers will most likely give you many surprises. For example the wellknown IEC connectors belonging to the cables between wall-outlet and the receiving-equipment (TV, videorecorder) appeared to be quite critical. That was already the case during our experiments in which we did not even use the cheapest connectors available for that purpose!

Measurements

Measuring digital signals is a completely different job than measuring analog signals. New measuring methods and - equipment are required. People with sufficient knowledge in this area are also needed. None of these things is available on a sufficient level at the moment. A number of measuring methods still have to be developed for this subject, where we hope to know at least what exactly is worthwhile to measure, for the digitized baseband signal as well as the HF modulated signal. Measuring equipment for the daily work in the field, is not at all available at reasonable prices at present.

The people responsible for the technical service of the CATV systems itself are facing an interesting future. However we can already determine, that the basic level of knowledge of these people must certainly be increased. Especially the people working in the head ends and so on have to be typologised as highly educated.

New applications

It will take some time of course before all the televisionprograms on satellite, cable and/or terrestrial will be transmitted in a digital manner. It is still quite uncertain whether we will switch off the last PAL TV transmitter in The Netherlands before the year 2015. However it is interesting to look ahead at the situation of the infrastructures as it could be within a decennium. CASEMA did so of course for her CATV systems. Resulting in an overview of the channel allocation on CATV systems that can be reached in an evolutionairy manner. We have given it the working title C60. Naturally the number of channels is much less than the possible number of programs. In the C60 system there are still 37 PAL TV channels/programs and in addition to that 25 DVB channels. The total number of programs that can be distributed depends on the applications for the DVB channels, but CASEMA estimates that on average 150 - 250 programs will be distributed simultaneously.

Among them there will be completely different types of programs than at present. A possibility in the future could be e.g. a channel that distributes the pictures of the individual camera's installed around a soccer field in parallel. That enables the viewer to select the camera that in is his opinion offers the best picture. Perhaps there will also be a 1 minute delayed picture available, enabling you to see a replay of the winning goal at your fingertips with your remote control! Another application is the ultimate form of NVOD (near video on demand): a particular movie film starts every -let us say- 5 minutes, divided over a number of channels. With such a solution there will be no further need to wait two hours or so for the end of the film when it restarts. Special programs for minorities can also be an attractive market. One should think about e.g. programs for (medical, scientific etc.) education. A number of cable-magazines instead of that current single one on most CATV systems is also a lucrative option. Most of those programs can be made with still pictures and sound and will require a very modest capacity of merely a part of a DVB channel and for that reason the costs of transmisson will be low.

Conclusions

Digital videotransmission via CATV networks offers enormous possibilities. Thanks to the growing capacity of the infrastructure and the reducing costs per hertz bandwidth an increasing number of new services can be realised. More services will lead to a better return on investments in the infrastructure. The relation between satellite and cable will be increasingly determined by competition as well as synergy.

In the head-ends of the CATV systems considerable investments will be necessary to generate a product that is tailormade for the cable-subscribers out of the -doubtlessly- extensive range of programs from satellites and other sources.

We must pay close attention to the technical quality of DVB signals on the cable television networks and the in-house systems. New measuring equipment and better trained service technicians are the primary preconditions for that. The public now imagine new video services as they will be in the digital era. Hopefully they will be able to enjoy all these new systems and techniques at reasonable costs.

@ 1995 N.V. CASEMA Technical Department Ir. F.J.W. van Let P.O. Box 670 2280 AR Rijswijk The Netherlands tel.: +31 (0)15 569284 fax: +31 (0)15 145530

Voordracht gehouden tijdens de 431e werkvergadering

STANDARDISATION IN THE DVB OF CONDITIONAL ACCES SYSTEMS FOR PAY TV

D. van Schooneveld
Philips Research Laboratories, Eindhoven, The Netherlands

Introduction

The topic of this paper is Pay TV as it is today and as it is envisaged in the near future. Special attention is given to the standardisation of Conditional Access systems for Pay TV in the DVB (Digital Video Broadcasting). Philips firmly believes in standardisation of these systems and aims at having the highest hardware commonality possible. This will reduce the price of the box which we feel is essential for deep market penetration and also in the interest of the consumer. The standardisation is all the more important since we are on the verge of the new digital transmission era, in which the consumer will be confronted with a wide range of new services.

Where are we today?

A popular definition of Pay TV broadcasting is that it is a form of television broadcasting where the service provider is in direct control of the customer's access to the service. The advantages for the provider of having direct control are twofold. Firstly, it guarantees revenues because customers who do not pay their bill can be denied access. Secondly, since the provider knows exactly how many customers are using his service, copyright negotiations with content owners are greatly simplified. The working of a Pay TV system is now briefly explained and at the same time, some often used terminology is introduced.

It is easiest to understand the system from the receiver side, where the system usually consists of a *set-top box* (also referred to as the *decoder box*), some detachable module such as a *smart card*, and the television set itself. The system will only work if the card is plugged into the box. The system works roughly as follows. Before transmission, the normal, or *clear* signal is *scrambled*. This means that it is distorted in such a manner that normal viewing of the service becomes impossible. At the receiver side, descrambling is carried out to convert the scrambled signal back to the clear signal. The clear signal is then fed into the television set. The *descrambler* itself is a piece of hardware located inside the set-top box, but its control is carried out by the *conditional access system*, located on the smart card. The conditional access system only allows the descrambler to operate if the customer is authorized to watch the service.

Authorization is done by sending a safely encrypted authorization message (*Entitlement Management Message, EMM*) to the customer which, after reception and decryption, is stored on the card. The authorization messages are usually sent well in advance of the service that they authorize. A second type of (encrypted) messages describes each service, and these messages *(Entitlement Control Messages, ECM's)* are sent together with the actual service data. Upon receiving an ECM, the smart card checks if the customer is authorized to receive the service, and only if this is the case, the descrambler is turned on by issuing of the appropriate key. These keys are often conveyed in encrypted form in the ECM's themselves.

On the transmission side, the system usually comprises several large subsystems that may be physically located in different buildings. A typical setup consists of the following equipment. First there is the equipment for playback and scheduling, usually located in the studio. There is also a computer based *Subscriber Management System, (SMS)*, needed to do all the administration and billing. The SMS includes an order entry system that passes the orders on to the *Subscriber Authorization System, (SAS)*. The SAS on its turn converts the orders into EMM's, which are then injected into the service. Other equipment necessary include a generator and an injector for the ECM's, and several control systems. All in all, the equipment on the transmission side is quite extensive and it is usually highly tuned to a specific conditional access system.

Over the past decade, several conditional access systems have been introduced. The following are only a few examples:

- One of the best known is VideoCrypt, developed jointly by News Datacom and Thomson Consumer Electronics. BSkyB, subsidiary of News International, which also owns News Datacom, makes extensive use of VideoCrypt for the wellknown transmissions of Sky Sports, Sky Movies, etcetera. The subscriber database of BSkyB contains over 2 million installed decoder boxes in the UK.
- A second well known system is Syster, developed by Canal+. The system is used by Canal+ on SECAM and PAL signals for terrestrial broadcasting. Canal+ has more than 4 million subscribers in their base, mainly in France. As another user, we name Premiere in Germany. Their box is one of the few that uses a special plastic key instead of a smart card.
- A third is Eurocrypt, developed by France Telecom, Bull and Philips. It is used on D2MAC signals for example by TV1000 and TV3 in Scandinavia.

All these systems are very sophisticated; they make use of various addressing schemes and allow different forms of authorization. The companies that have pioneered these systems have spent huge amounts of time and effort on the research, development and manufacturing of the boxes, Subscriber Management Systems, Subscriber Authorization Systems, message injectors and not in the least in the setting up of the subscriber base itself. Their investments should not be underestimated.

The incompatibility of the systems however poses a problem to the customer; if he wishes to receive another service, he has to rent or buy an additional decoder box and smart card. Especially if the other service is new on the market, it is really not likely that the customer is willing to do that. A solution to the problem could be to let this new service be accessed through the already installed boxes, which means that the two service providers would have to come to some mutual agreement. Recent examples of this can be found in the UK where BSkyB has reached such agreements with TV Asia, UK Gold and many more.

Standardisation of Pay TV systems in the DVB

Digital Video Broadcasting

Television broadcasting is nowadays rapidly progressing towards the digital era. Soon, transmission of television signals will be done in a fully digital format, using audio and video compression techniques (MPEG-2) that make use of the frequency spectrum in a far more efficient manner than the analog techniques currently in use. As a direct result, the costs for using a transponder (per service) will drop, new broadcasters are likely to emerge quickly and the number of services offered to the consumer will explode. This leads to one of the reasons for the existence of Pay TV: it is expected that the audience will choose for higher quality channels that offer services better suited to their needs, and also that the audience is willing to pay extra for these services.

Standardisation of Pay TV systems is formally discussed in the European Digital Video Broadcasting (DVB) Project. The DVB Project is an ensemble of committees formed by representatives of network providers, satellite operators, broadcasting, telecommunications and manufacturing companies from all over Europe.

The aim of the DVB Project is to come to standards on transmission of digital video data through the air (terrestrial), through cable and by satellite. Issues on the table are for instance modulation and channel coding, service information, multiplexing and also conditional access. The ultimate goal is to create a business environment where consumers are reasonably well protected and where various providers of services and equipment can compete on fair terms and where new actors can enter without excessive thresholds.

The DVB Project started in 1991, when a so called Technical Module was established. Since then, the Project has steadily grown in size; over 150 companies that wish to participate in the discussions have today signed the Memorandum Of Understanding.

In the DVB Project, two groups focus on scrambling and conditional access. Both groups study the feasibility of a scheme where the systems share as much commonality as possible. The first, a commercially driven group called the Ad-Hoc Group on Conditional Access *(Ad-Hoc group on CA)*, considers complete conditional access systems, including scramblers/descramblers and periferal equipment at the transmission side. The second group, concerned with the more technical issues and called the Conditional Access Specialist Group (CASG), focusses on the decoder box and transcoding issues.

Standardisation

Although formal standardisation may be problematic, it is generally agreed within the DVB Project that standards are essential for digital television broadcasting. This consensus does however not extend to conditional access and scrambling.

Among the members of the commercially driven Ad-Hoc group, there are various points of view. There are the broadcasters, who may be classified either as "established pioneers" or as "newcomers". On the one hand, the pioneers would like make sure that their investment in their operation is worth every penny and will accept standardisation only up to a certain point. On the other hand, the newcomers, consisting mainly of network operators and the smaller broadcasters, would like to cooperate, since they know that not many people would buy a decoder box to watch only one or two channels. Therefore, complete standardisation of the box would really be the solution for them. Finally there are the equipment manufacturers, such as Philips, who wish to standardise to come to cheap mass production.

The more technically oriented Specialist Group has looked at several scenarios involving different combinations of scramblers and conditional access systems. It was quickly agreed by all parties that the decoder box should at least contain some standardised descrambler, because nobody wants the consumer to be forced to buy multiple boxes. It appeared however that, due to export regulations, issues related to intellectual property rights and a general lack of suitable candidates, it was not so easy to find this standard. In the end, the group nominated a small subgroup of cryptologic experts that came up with a solution after four months. Various scenarios for the descramblers and the conditional access systems were then analyzed and the result was two proposals that were put to the *Steering Board* of the DVB (more or less comparable with the Board of Directors of a company).

The first proposal is now commonly known under the name of *Simulcrypt*. This is more or less what has happened in the UK with TV Asia and UK Gold (see a previous section); the "newcomer" that wishes to access the "pioneers" set of installed decoder bozes, first has to come to a business agreement with the pioneer who then ensures that his CA system also provides the newcomers service. This means that the newcomer has to deliver the scrambling keys and the subscriber information to the pioneer, who then returns the corresponding ECM's and EMM's and makes sure that the smart card is capable of the newcomers' service. The proposal is heavily backed by the pioneers (BSkyB, Canal+, Filmnet and their respective CA system suppliers.)

The main problem for those who oppose this proposal is of course its necessity to sign an agreement. Therefore it was decided to draft a standard agreement to ensure that business takes place under fair and reasonable terms. This standard agreement is now known as the *Code of Conduct* and still has to be finalized.

The second proposal is called *Multicrypt* but is better known as the *Common Interface*. The Common Interface today consists of a detailed draft specification of a standard interface between the decoder box and a detachable module. Since the interface would have to fit all incompatible CA systems in use today, the interface was chosen at the MPEG-2 Transport Layer. This means that scrambled MPEG-2 data goes across the interface into the module and descrambled (thus clear) data is returned, both at a rate of up to 50 Mbit/sec. The box itself is completely standardised and does not contain the descrambler chip any more; this chip is now inside the detachable module. Since there are no suitable alternatives, the detachable module was chosen to be a PCMCIA card, currently in use for computer applications. For the Pay TV application, this card may contain more than the descrambler chip and the CA system, for instance software for an electronic programme guide. The cost of the card and that of its impact on the decoder box software is not yet clear.

The Common Interface is backed by the newcomers, who see this as a fundamentally good and secure solution to the standardisation problem; it allows them to cooperate and still use their own *proprietary CA system*.

Since it appeared to be impossible to come to a compromise, the Steering Board in the fall of 1994 accepted a package that Consisted of several elements, among which recommendations for anti-piracy legislation, transcoding issues and both the mentioned proposals: Simulcrypt and the Code of Conduct as well as the Common Interface, but neither of them mandatory. This means that a broadcaster who wishes to be DVB compliant still has both options; he can either use a proprietary system (with the standardised scrambler) and let any newcomer enter his market under the rules laid down in the Code of Conduct, or he can use the Common Interface.

The near future

Now that neither of the two options are mandatory, it is expected that both solutions will emerge and co-exsist in the near future. The CA package proposed by the Steering Board still has to be accepted by the E.C. Council and it is expected that those in favour of the Common Interface will try to push the proposal to become mandatory. A disadvantage for them is that the proposal still has to be finalised and there appears to be some work left. Especially the choice of the detachable module (PCMCIA) is not clear and several parties argue about its price. If the Common Interface is finalised soon and if it indeed becomes a mandatory standard, then it is expected that the established broadcasters, who are in favour of the Simulcrypt proposal, will put their orders for new digital systems on hold for quite a while.

Voordracht gehouden tijdens de 431e werkvergadering

UIT HET NERG

Verslag van de Algemene Ledenvergadering van het NERG d.d. 29 maart 1995

Aanwezigen:

Leden:

de heren ir. G.A. Joosten, prof.ir. L. Krul, dr.ir. W. Herstel, ir. R.A. Kasper, ir. A.J.G. Dorgelo, ir. D.A. van der Meij, ir. J. van Egmond, ing. J.F. Deckwitz, ir. P.J.T. Bruinsma, Ph.J. Huls, ir. C.D. de Haan, ir. H.M. Schuit, ir. M.W.A. Groenewegen, ir. R.J. Kopmeiners, ir. A.A.J. Otten, P.F. Maartense, dipl.ing. J. Hekner, ing. J. Maas, ing. H.C. Milius, ir. J.B.F. Tasche, A. van der Zwan (administrateur), ir. W.F.M. Groenewegen, prof.dr.ir. J. Davidse, dr. M. Jeuken, J.P.A. Lamb, R.W. Budding, ir. L. Esser, ir. J. van Duuren, ir. D.W. Rollema.

Bestuur:

de heren ir. P.G.M. Baltus (aftredend), ir. W. van der Bijl, prof.dr.ir. W.M.G. van Bokhoven, dr.ir. R.C. den Dulk(aftredend), prof.ir. J.H. Geels (voorzitter), ir. G.J. de Groot (verslag, aantredend), ir. O.B.P. Rikkert de Koe, ir.P.R.J.M. Smits, ing. A.A. Spanjersberg (aantredend), dr .ir .drs. E.F. Stikvoort (aantredend).

1 Opening

Om 10.00 uur opent de voorzitter de vergadering. Hij vraagt enige ogenblikken stilte wegens het overlijden van dhr. P.J.M. Geenen, ir. M.G.A. Haalebos, prof.ir. C. Rodenburg en ir. H.G.W. van den Steen.

2 Verslag van de vorige Algemene Ledenvergadering d.d. 30 maart 1994.

Pagina 1 De naam van dipl.ing. J. Hekner is fout gespeld in de lijst van aanwezigen (was Henker).

Pagina 2 Naar aanleiding van de samenwerking met AES (paragraaf 1.3) wordt opgemerkt dat deze samenwerking in 1994 minder intensief was.

Pagina 3 Naar aanleiding van de suggestie in paragraaf 3.4 om contact met de Poolse Vereniging voor Ingenieurs te zoeken, wordt de vraag gesteld wat op dit gebied is gebeurd. De voorzitter antwoordt dat hier nog niet veel energie in is gestoken. Er zijn wel initiële contacten geweest, maar daarna is niets meer vernomen.

De tekst "... waarmee het bestuur decharge wordt verleend." in paragraaf 3.5 moet vervangen worden door: "... wat aanleiding geeft tot decharge van de bestuursleden."

De tekst "De heer Tasche wenst na een bestuursperiode van vijf jaar af te treden." moet vervangen worden door: "De heer Tasche treedt statutair af na een bestuursperiode van zes jaar (waarvan vijf als voorzitter)."

Na deze wijzigingen en opmerkingen wordt het verslag vastgesteld.

3 Jubileum: 75 jaar NERG.

De voorzitter deelt mee dat in verband met de financiële situatie van het NERG het oorspronkelijk budget voor het jubileum gehalveerd is. Desondanks is er een aantrekkelijk jubileumprogramma samengesteld. Het 75-jarig jubileum start 23 mei a.s. met een bijzondere Algemene Ledenvergadering tijdens een bezoek aan de Amercentrale. Er zullen in het jublileumjaar twee extra nummers van het Tijdschrift verschijnen (een index van alle artikelen van de afgelopen 75 jaar en een speciale uitgave). De werkvergaderingen zullen gedurende dit jaar in het teken staan van het doorlopen van het frequentiespectrum. Bovendien wordt een CD-ROM uitgegeven. Op een vraag uit de vergadering deelt de voorzitter mee dat leden die niet in het bezit zijn van een PC met CD-ROM speler deze ondermeer in openbare bibliotheken kunnen afspelen.

4 Jaarverslag van het NERG en aanverwante organisaties over het jaar 1994.

Het verslag wordt pagina voor pagina doorgenomen. De volgende vragen worden gesteld:

- Naar aanleiding van paragraaf 2.3 (Onderwijs/onderwijscommissie) wordt de vraag gesteld welke invloed het NERG heeft op de normen voor de Rens & Rens Hoger en Middelbaar Elektronica examens. De voorzitter van de Onderwijscommissie, prof. van Bokhoven, antwoordt dat binnenkort hieromtrent overleg gevoerd wordt met de gecommitteerden.
- Naar aanleiding van paragraaf 3.1 (SVEN-fonds NERG) wordt de vraag gesteld hoe vaak gebruik is gemaakt van de subsidieregeling voor het presenteren van publikaties op symposia. Prof. van Bokhoven antwoordt dat dit één maal gebeurd is.
  Verder deelt hij na een vraag mee dat in 1994 een aanzet gemaakt is voor het opzetten van faciliteiten voor het geven van onderwijs op afstand. Als medium zal hiervoor Internet gebruikt worden, met behulp van een server op de TUE. Het zal nog minimaal een jaar duren voor dat dit operationeel is.
- Naar aanleiding van paragraaf 4.1.1 (NERG-exploitatierekening 1994) wordt de vraag gesteld uit welke componenten de administratiekosten bestaan. De voorzitter antwoordt dat de administrateur drie hoofdtaken heeft: ondersteuning van de penningmeester (leden-, donateurs- en abonnementenadministratie), ondersteuning van de secretaris (postbehandeling en archivering) en ondersteuning van de programmamanager en hoofdredacteur (de verzendlijsten voor uitnodigingen en de inboeking van deelnemers aan werkvergaderingen en de verzendlijst van het Tijdschrift). Door het uitbesteden van enkele deeltaken is de verwachting dat op deze post bezuinigd kan worden.
- Naar aanleiding van paragraaf 4.1.2 (Balans NERG per 31 december 1994) wordt de vraag gesteld of het Onderwijsfonds niet meer bestaat. De voorzitter licht toe dat dit in het vermogen van het NERG is onder gebracht, aanleiding was het wegvallen van de bestaansgrond nadat het SVEN-fonds bij het NERG was gekomen, zoals reeds in een vorig jaar verslag is vermeld.
- Het vermogensverlies is volledig ten laste van het Fonds Bijzondere Activiteiten (FBAC) geboekt.
- De heer Tasche vraagt op welke termijn de vermogensvermindering tot problemen leidt. De voorzitter antwoordt dat dit binnen vijf jaar zal zijn als bijzondere baten uitblijven.

De voorzitter leest het verslag voor van de kascontrole commissie, waarin geconcludeerd wordt dat de financiële activa aanwezig zijn en dat boekhouding op correcte wijze gevoerd is. Hierna verleent de vergadering de bestuursleden decharge van het in 1994 gevoerde beleid. De vergadering blijkt vol lof te zijn over de opzet en inhoud van de verslaggeving.

5 Jaarplan van het NERG en aanverwante organisaties voor het jaar 1995.

Het jaarplan wordt door de voorzitter gepresenteerd. De vraag wordt gesteld of er een structurele vorm van samenwerking met zusterverenigingen is gepland op het gebied van de "rode draad" in de werkvergaderingen. Het bestuur antwoordt dat de plannen voor deze cyclus werkvergaderingen gepresenteerd zijn tijdens het gezamenlijk overleg. Hoewel nog geen concrete plannen zijn gemaakt, wordt verwacht dat voor sommige werkvergaderingen samenwerking mogelijk is.

6 Perspectief van het financieel beleid 1996 en volgende jaren.

De voorzitter geeft een toelichting op de verdeling van baten en lasten en op welke wijze het structurele tekort bestreden kan worden. Aan de inkomstenkant kan de situatie verbeterd worden door een contributieverhoging naar *f* 75,- en nog meer leden te werven. Gestreefd wordt naar ten minste 35 nieuwe leden per jaar. Als daarnaast een bedrag van *f* 8000,- bespaard wordt op de administratie- en bestuurskosten (nu te samen 40% van de uitgaven), kan bij een inflatie van 4% gedurende lange tijd een gezonde financiële situatie gehandhaafd worden. De vraag wordt gesteld waarom achterstallige contributies niet geïnd kunnen worden. Dit is vaak moeilijk, bijvoorbeeld omdat leden zoek raken door verhuizing of omdat kennismakers ongewild als lid worden ingeschreven. Opgemerkt wordt dat enig optimisme op zijn plaats zou zijn omdat na zeven magere jaren, weer zeven vette jaren zullen volgen. De voorzitter antwoordt dat het bestuur geen bijzondere baten in het verschiet ziet. De vergadering spreekt zijn waardering uit voor de plannen van het bestuur om orde op zaken te stellen met betrekking tot het financiële beleid.

7 Voorstellen tot contributieverhoging ingaande het jaar 1996.

De vergadering keurt unaniem het voorstel goed om de contributie voor gewone leden te verhogen van *f* 60,- naar *f* 75,- per jaar en die voor juniorleden van *f* 30,- naar *f* 39,- per jaar. De 50% korting voor jonge leden wordt afgeschaft. Wanneer men een machtiging tot automatische incasso aan het NERG verleent, wordt een korting van *f* 3,- gegeven wegens besparingen op de administratiekosten. De penningmeester merkt op dat deze korting niet van toepassing is op automatische periodieke overschrijvingen. Uit de vergadering wordt opgemerkt dat in het recente verleden leden juist aangespoord zijn de contributie te voldoen via een automatische periodieke overschrijving. Moeten zij nu deze opdracht omzetten in een machtiging tot automatische incasso om in aanmerking te komen voor de korting? Het bestuur zegt toe hier soepel mee om te gaan.

8 Voorstel tot wijziging van artikel 2 van het Huishoudelijk Reglement

Naar aanleiding van het voorgaande agendapunt keurt de vergadering met algemene stemmen de volgende tekst voor artikel 2 van het Huishoudelijk Reglement goed:

CONTRIBUTIE, DONATIE

Artikel 2

1. De contributie dient telkenjare voor 1 februari van het verenigingsjaar, waarover zij verschuldigd is, te worden voldaan door storting of overschrijving op een door de penningmeester bekend te maken rekening. Op de verschuldigde contributie wordt een korting wegens kostenbesparing verleend aan alle leden die een machtiging tot automatische incasso aan het bestuur hebben verleend.
2. Zij, die in de loop van enig jaar als gewoon lid of juniorlid worden toegelaten, zijn over dat jaar een contributie naar gelang van de tijd verschuldigd, te berekenen als het produkt van het aantal na de datum van toelating tot het lidmaatschap nog niet verstreken gehele maanden en het twaalfde deel van de voor het lidmaatschap vastgestelde contributie; het aldus berekende bedrag wordt naar beneden afgerond op gehele guldens.
3. In bijzondere gevallen kan het bestuur beslissen dat een gewoon lid of juniorlid wegens persoonlijke omstandigheden geheel of gedeeltelijk van contributiebetaling zal zijn vrijgesteld. Deze beslissing geldt telkens voor één verenigingsjaar.
4. Degenen, die door het bestuur met het oog op lidmaatschap worden uitgenodigd kennis te maken met de vereniging, ontvangen gedurende ten hoogste twaalf maanden om niet de uitnodigingen voor de werkvergaderingen en de afleveringen van Het Tijdschrift.
5. Donateurs betalen een jaarlijkse bijdrage ter grootte van ten minste twee maal de contributie van een gewoon lid, te voldoen voor 1 april van het betreffende verenigingsjaar.

Het gewijzigde artikel zal in het Tijdschrift worden gepubliceerd.

9 Verkiezingen

9.1 De voorzitter deelt mee dat de volgende bestuursleden aftredend zijn:

- Statutair aftredend en herkiesbaar zijn: ir. P.R.J.M. Smits, ir. P.K. Tilburgs en ir. C.Th. Koole. De heer Tilburgs kan zich niet langer beschikbaar stellen.
- Statutair aftredend en niet herkiesbaar zijn: dr.ir. R.C.den Dulk en ir. P.P.M. van der Zalm.
- Tussentijds aftredend op eigen verzoek: ir. P.G.M. Baltus

De vergadering gaat bij acclamatie accoord met de (her)benoeming van de volgende personen in het bestuur:

- Herbenoemd in het bestuur worden ir. P.R.J.M. Smits (vice-voorzitter) en ir. C.Th. Koole (lid).
- Benoemd worden de nieuwe bestuursleden ir. G.J. de Groot (secretaris), ing. A.A. Spanjersberg (hoofdredacteur), dr.ir.drs. E.F. Stikvoort (lid) en dr.ir. A.P.M. Zwamborn (lid).
- De heren D.J. Benne en ir. P.B. Hesdahl worden gekozen in de kascommissie voor het boekjaar 1995 met als plaatsvervangende leden ir. A. van Schelven en ir. Chr.H.M.Clemens.
- De ALV herbenoemt ir. C.D. de Haan (statutair aftredend en herkiesbaar) in de Ballotagecommissie.

10 EUREL

De voorzitter geeft een toelichting met betrekking tot het voornemen van het bestuur om het lidmaatschap van EUREL te beëindigen. De voornaamste reden is dat EUREL te veel is weggedreven van de oorspronkelijke motieven voor de samenwerking van verenigingen van elektrotechnisch ingenieurs. Het NERG heeft niet de organisatie en de middelen om aan de huidige EUREL activiteiten volwaardig mee te doen. Verder is het bestuur van mening dat het NERG een onevenredig hoge contributie aan EUREL moet betalen. Dit wordt veroorzaakt door het contributiestelsel van EUREL waardoor kleine verenigingen, zoals het NERG, per lid relatief veel meer betalen dan grote verenigingen. Als voorbeeld wordt de vergelijking met IEE uit Engeland gemaakt. Per lid betaalt het NERG tenminste 35 maal zoveel aan contributie. Na opening van het Brussels kantoor van EUREL werd de contributie verviervoudigd. Sindsdien betaalt het NERG uit protest de oude contributie. Aan deze situatie zou per 1 januari 1996 een eind komen. EUREL stelt voor dat het NERG samen met KIvI een nationale koepel zou moeten vormen om de kosten te verminderen. Het KIvI overweegt echter de "enhanced service" in plaats van de " basic service". Samenwerking zou voor het NERG dus niet tot een kostenreductie leiden. Na deze uitleg gaat de vergadering accoord

met het plan het lidmaatschap van EUREL op te zeggen.

11 Rondvraag.

Er wordt geen gebruik gemaakt van de rondvraag.

12 Sluiting

om uiterlijk 10.45 uur. Met excuses voor de onvoorziene langere duur van deze vruchtbare vergadering sluit de voorzitter om 11.15 uur.

LEDENMUTATIES

Voorgestelde leden:

B. van Keken	Gounodstraat 13	3816 WH Amersfoort
ing. J.B.M. Kollenbrander	Leeuwenburg 45	3401 HV IJsselstein
ir. H.A.H. Otten	Oude Delft 29 D	2611 BB Delft
ing. W.A.M. Snijders	De Zicht 15	5502 HV Veldhoven
prof.ir. J. de Stigter	Irenestraat 5	2731 GE Benthuizen

Nieuwe Leden:

V.B. Grundlehner	Staringstraat 22	5615 HD Eindhoven
X. Lin	Backenhagen 7	5655 KZ Eindhoven
ir. M. Lugthart	Kempering 928	1104 KE Amsterdam
ir. R. Middelkoop	Eisensteinstrook 8	2726 SX Zoetermeer
ir. G. Pappot	Middenweg 94	1394 AM Nederhorst den Berg
ir. K.J. van Staalduinen	Ruwaarddwarsstraat 3- B	3023 PZ Rotterdam
ir. P.P. Vervoort	Rietmolenstraat 50-10	7512 XW Enschede
ir. P.J. de Vrijer	Calsstraat 20	8015 BK Zwolle

Nieuwe adressen van leden:

ir. J.H. Dijk	Johan Westerveldstraat 16	1318 CC Almere
ir. E.A. de Boer	Bosboom Toussaintsingel 58	2741 AN Waddinxveen
ir. T. Borst	Spadestraat 2	4818 EL Breda
ir. J.C. Dito	Landlustlaan 108	2265 EK Leidschendam
ir. H. van Glabbeek	Zevenhuizen 3	5595 XE Leende
ir. G.C. Groenendaal	Julianalaan 4	5062 JM Oisterwijk
ir. R.J. Kopmeiners	A. van Schendelstraat 703	3511 MZ Utrecht
ir. F.J.M. van Kuppeveld	Westvest 26	2611 AZ Delft
ir. M.L. Lugthart	De Pruikenmaker 9	5506 CT Veldhoven
prof.ir. O.W. Memelink	Buitenweg 211	3602 XB Maarssen
ir. G.A. Niekolaas	Melis Stokelaan 143	2543 GJ Den Haag
ir. J.C. van der Plaats	Reigerskamp 801	3707 JW Maarssen
ir. H.J. Simons	Lange Kerkdam 68	2242 BX Wassenaar
ir. F.J. Sluijs	Ulenpas 55 A	5655 JB Eindhoven
ir. E.R. Smit	Kanaal OZ 9-63	9419 TJ Dryber
dr.ir.drs. E. Stikvoort	Pisanostraat 138	5623 CE Eindhoven
ir. E. Stringer	Stuyvesantstraat 8	2023 KN Haarlem
dr.ir. R.N.J. Veldhuis	Staringstraat 15	5615 HC Eindhoven
ir. R.M. Wiegers	Wisentstraat 20	6532 AP Nijmegen
ir. F. Zelders	Plataanhout 42	2719 KK Zoetermeer

IEEE
Benelux Section

Symposium Announcement

IEEE Third Symposium on Communications and Vehicular Technology in the Benelux

October, 25-26, 1995

SCVT '95 is the third Symposium organized by the IEEE Benelux joint Chapter on Communications and Vehicular Technology.

Chapter committee: Prof. R. Prasad, Chairman, TU Delft, The Netherlands. Prof. A. Laloux, Vice-Chairman. UCL, Belgium, J. Farserotu, Secretary, K. Smit, Treasurer, L. Vandendorpe, Secretary cum Treasurer, S. Baggen, Inter society and European, activity Editor, Chapter Letter, M. Durvaux, Awards and nominations, G.J.M. Janssen, Student relations, A.S. Krishnakumar, Workshop coordination, M. Moeneclaey, Symposium coordination chair, P.F.M. Smulders, membership chair.

Program: Please find enclosed the advance Program of the Symposium.

Language: The Symposium language for all presentations is English. No translation facilities will be provided. The contributions will be published in proceedings available during the Symposium at the registration desk.

Symposium dinner: Symposium participants and their companions are invited to the Symposium dinner which will be held on Thursday, October 25.

Paper awards: A best master student paper award as well as a best paper award for Ph.D. candidates will be issued at the end of the Symposium.

Registration: Please find enclosed a Registration Form of the Symposium which should be received before October 4, 1995. It is also possible to pay cash (Dutch currency) or by credit card at the registration desk.

Venue: The Symposium will be held in the Dorgelo/v. Trier Room in the Management Building of the University of Technology in Eindhoven. The University buildings are situated near the railway station. Hence, you can reach the university quite easily by train. You can get direct flights to Eindhoven Airport from various cities in Europe. Information of how to get to the meeting place will be send to you after the symposium registration form has been received.

Hotels: Please find enclosed a Hotel Reservation Form. Block reservations have been made in these hotels. When you want to benefit from the reduced rates please send in the Hotel Reservation Form before October 4, 1995.

Information: For any additional information please contact Peter Smulders, tel: 31 40 473662, fax: 31 40 455197, email: P.F.M.Smulders@ele.tue.nl

WET PERSOONSREGISTRATIE

Deze wet stelt regels ter bescherming van de persoonlijke levenssfeer in verband met persoonsregistraties. Daaronder wordt in de wet verstaan: een samenhangende verzameling van op verschillende personen betrekking hebbende persoonsgegevens, die langs geautomatiseerde weg wordt gevoerd of met het oog op een doeltreffende raadpleging systematisch is aangelegd.

Om het werk goed te kunnen doen is het bestuur van het NERG zowel houder als bewerker van zo'n persoonsregistratie, waardoor het moet voldoen aan de in de wet gestelde gedragsregels:
Een persoonsregistratie mag slechts worden aangelegd voor een bepaald doel waartoe het belang van de houder redelijkerwijs aanleiding geeft. Uit een persoonsregistratie worden slechts gegevens aan een derde verstrekt voor zover zulks voortvloeit uit het doel van de registratie, wordt vereist ingevolge een wettelijk voorschrift of geschiedt met toestemming van de geregistreerde.
De houder is verplicht mededeling te doen aan de geregistreerden over het doel, het beleid en de inhoud van de registraties.

Dit bericht dient om de leden te laten weten hoe het met de persoonsregistratie binnen het NERG is gesteld.

Het werk van het NERG houdt in dat van de leden en kennismakers in ons kantoor te Leidschendam voor zo ver van toepassing worden geregistreerd:

- naam, voorletters, titulatuur, adres, postcode, woonplaats;
- kwalificatie voor lidmaatschap, datum toelating tot lidmaatschap, categorie van lidmaatschap, datum einde lidmaatschap;
- bankrekening, gegevens inzake de contributievoldoening;
- vervulde functies binnen het NERG en de periode daarvan.

Deze registratie is van belang voor:

- het bewaken van de status van het lidmaatschap;
- het periodiek opstellen van een ledenlijst;
- de inning van contributies;
- het opstellen van de verzendlijst van de uitnodiging voor ledenvergaderingen en werkvergaderingen;
- het opstellen van deelnemerslijsten van deze werkvergaderingen;
- het opstellen van de verzendlijst van Het Tijdschrift;
- de correspondentie met de leden.

Van bestuursleden worden in verband met de wettelijk verplichte inschrijving in het Verenigingenregister ook de voornamen en de geboortedatum geregistreerd.

De drukker van de uitnodigingen en Het Tijdschrift verzorgt thans ook de verzending. De verzendlijsten worden aan hem ter hand gesteld onder voorwaarde dat hij daarvan slechts gebruik mag maken voor de verzending en dat elk ander gebruik verboden is.

De deelnemerslijst van een werkvergadering wordt aan de gastheer gegeven als een opgave van namen nodig is, b.v. met het oog op zijn regels betreffende ontvangst van bezoekers. Daarbij geldt dat het bestuur elk ander gebruik verbiedt.
De periodiek gepubliceerde ledenlijst vervult een belangrijke functie bij het elkaar vinden van de leden in het kader van de doelstellingen van het NERG. Het bestuur stelt de ledenlijst niet voor een ander gebruik (b.v. commercieel) beschikbaar. Indien een lid schriftelijk bezwaar maakt tegen zijn vermelding in deze ledenlijst, zal het bestuur daarvan aantekening maken en hieraan bij de eerstvolgende uitgave gevolg geven.

De donateurs en abonnees zijn ook opgenomen in een bestand. In het algemeen zijn dat geen natuurlijke personen, dus vallen ze niet onder de wet persoonsregistraties. Het NERG gaat daarmee niettemin op analoge wijze om als voor de leden is geschetst.

JHG

Cursusaankondigingen

PATO

- Bedrijfszekerheids-gegevens 3 en 4 oktober 1995 in Arnhem
- Bedrijfskundige aspecten van hergebruik 26 en 27 oktober 1995 in Oisterwijk
- Optische Communicatie 1, 2 en 3 november 1995 aan de TU Eindhoven
- Laagspannings- en laagvermogens analoge IC-techniek 2 en 3 november 1995 in Delft
- Geavanceerde meettechnieken voor plaats- en krachtgrootheden 6, 13 en 20 november 1995 in Utrecht
- Nutswereld in beweging 8 en 9 november 1995 in Utrecht
- RAM (Reliability, Availability, Maintainability) Optimalisatie en Ontwerpbeoordeling met FMECA (Failure, Modes, Effects, Criticality, Analysis) 8 en 9 november 1995 in Arnhem
- Integratie van bedrijfsketens 8, 9 en 10 november in Delft
- Elektro-magnetische Compatibiliteit 16-17, 23-24, 30 november en 1 december 1995 in Eindhoven
- Effectief ontwikkelen van nieuwe producten 21, 22 en 23 november in Oisterwijk
- Machine Vision 21, 22, 28 en 29 november 1995 in Eindhoven
- Taguchi Methoden 23, 30 november en 7 december 1995 in Eindhoven

Contactadres: Stichting PATO, Prinsessegracht 23
Postbus 30424, 2500 GK Den Haag
Tel: 070 3644957 Fax: 070 3562722

CEI-EUROPE

Advanced Technology Short Courses on Telecommunications and Related Technologies: oktober, 1995 Baveno/Stresa, Italy:

- Digital Receivers for Satellite and Mobile Communication, 2-5 oktober
- Applied RF Techniques I - Linear Circuits, 2-6 oktober
- Applied RF Techniques II - Nonlinear Circuits, 2-6 oktober
- Telecommunication Software Architecture and Design, 2-6 oktober
- Digital Design for High Speed Cicuits and Systems, 2-6 oktober
- Spread Spectrum/ CDMA, 2-6 oktober
- Cellular and Personal Communications Infrastructure, 9-11 oktober

oktober, 1995 Barcelona, Spanje:

- Speech and Channel Coding for Mobile Communication, 16-18 oktober
- Signalling for Telecommunications: Fixed and Mobile Networks, 16-19 oktober
- Far-Field, Compact and Near-Field Antenna Measurement Techniques, 16-19 oktober
- Modelling and Simulation of Communication Systems, 16-19 oktober
- VSAT Networks, 18-20 oktober
- Telecommunication Network Management, 23-26 oktober
- Error Correcting Codes and Trellis-Coded Modulation, 23-26 oktober
- High-Speed Data Communication over Wire-Pair Channels, 23-26 oktober
- Analog Circuit Design for Data Converters, 23-27 oktober
- Modern Digital Modulation Techniques, 23-27 oktober
- Satellite Communication Systems, 23-27 oktober
- Wireless Digital Communications, 23-27 oktober
- Personal Mobile Satellite Communications, 26-27 oktober

International Courses in Advanced Technology:

- Neural Networks and Pattern Recognition
- Digital Cellular & PCS Communications - The Radio Interface
- Analog Signal Processing & Related CMOS Circuit Design

13-17 november, Cambridge, UK

- High Frequency Analog Cicuit Design for Communication Systems
- Telecommunication Switching Systems
- Next Generation Networks - The Information Superhighways
- Enterprise Networking

20-24 november 1995, Cambridge, U.K.

Contactadres: Mrs. Tina Persson
CEI-Europe, Box 910, S-612 25 Finspong, Sweden
Tel: +46 122 17570 Fax: +46 122 14347

Conferentie aankondigingen

IEE - IEEE

International Conference on Clean Electronics Products and Technology
9-11 oktober 1995, Edinburgh International Conference Centre, Scotland

Contactadres: EE'95 Secretariat
Savoy Place
London WC2R 0BL UK
Tel: +44(0) 171 344 5477/5478
Fax: +44(0) 171 240 8830
Email: conference@iee.org.uk

AGARD (Advisory Group for Aerospace Research & Development):
Knowledge-based Functions in Aerospace Systems
6-7 november 1995 in Madrid, Spanje
9-10 november 1995 in Châtillon, Frankrijk
16-17 november 1995 in Moffet Field, California USA

Inlichtingen: Tel: (33.1)47.38.57.14 (of 16)
Fax: (33.1)47.38.57.99
Place des Nations
CH-1211 Genève

Call for papers

IEE

10th European Frequency and Time Forum
5-7 maart 1996, Brighton, UK
deadline for summaries: 16 oktober 1955

IEE Inmarsat

5th International Conference on Satellite Systems for Mobile Communications and Navigation
deadline for summaries: 16 oktober 1995

Inlichtingen: Tel: +44 171 344 5467 / 5478 / 5477
Fax: +44 171 240 8830
Email: conference@iee.org.uk

Tijdschrift van het Nederlands Elektronica- en Radiogenootschap

deel 60 nr.3 - 1995

Inhoud

Druk en DTP: Drukkerij Van Ek, Driebergen